Red Spider Meets a Rainworm: Conjunctive Query Finite Determinacy Is Undecidable

Red Spider Meets a Rainworm: Conjunctive Query Finite Determinacy Is Undecidable.

Tomasz Gogacz, Jerzy Marcinkowski,
Institute of Computer Science, University Of Wrocław,

***Abstract*—We solve a well known and long-standing open problem in database theory, proving that Conjunctive Query Finite Determinacy Problem is undecidable. The technique we use builds on the top of our Red Spider method which we developed in our paper [GM15] to show undecidability of the same problem in the "unrestricted case" – when database instances are allowed to be infinite. We also show a specific instance Q_0, $\mathcal{Q} = \{Q_1, Q_2, \ldots Q_k\}$ such that the set $\mathcal{Q}$ of CQs does not determine CQ Q_0 but finitely determines it. Finally, we claim that while Q_0 is finitely determined by $\mathcal{Q}$, there is no FO-rewriting of Q_0, with respect to $\mathcal{Q}$, and we outline a proof of this claim[1].**

I. Introduction

"Assume that a set of derived relations is available in a stored form. Given a query, can it be computed from the derived relations and, if so, how?" is the first sentence of [LY85]. Saying the same in today's language:

> The instance of the Conjunctive Query Finite Determinacy Problem (CQfDP) problem is a set of conjunctive queries $\mathcal{Q} = \{Q_1, \ldots Q_k\}$, and another such query Q_0. The question is whether $\mathcal{Q}$ determines Q_0, which means that for each two database instances (that is **finite** relational structures) $\mathbb{D}_1$ and $\mathbb{D}_2$ such that $Q(\mathbb{D}_1) = Q(\mathbb{D}_2)$ for each $Q \in \mathcal{Q}$, it also holds that $Q_0(\mathbb{D}_1) = Q_0(\mathbb{D}_2)$.

Answering queries using views appears in so many various contexts that it is indeed hard to imagine a more natural scenario in database theory. See for example [H01], or a recent thesis [F15] for a survey[2]. The contexts where such scenario appears include for example query optimization and caching [DPT99], or – to see more recent examples – [FG12] where the view update problem is studied and [FKN13] where the context are description logics. In all the examples we mentioned so far we "prefer" Q_0 to be determined by $\mathcal{Q}$. Yet another context, where it is "preferred" that Q_0 is not determined, is privacy: we would like to release some views of the database, but in a way that does not allow certain query to be computed.

Many variants of the problem were being considered, and the case we study, where both the views and the query are defined by conjunctive queries, and where the views we can see are "exact" is not the only possible (what we call "determinacy" is "losslessness under the exact view assumption" in the language of [CGLV00]). Let us just cite the most recent results: [NSV10] shows that the problem is decidable if each query from $\mathcal{Q}$ has only one free variable; in [A11] decidability is shown for $\mathcal{Q}$ and Q_0 being "path queries". This is generalized in [P11] to the the scenario where $\mathcal{Q}$ are path queries but Q_0 is any conjunctive query. The paper [NSV07] is the first to present a negative result. It was shown there, that the problem is undecidable if unions of conjunctive queries are allowed rather than CQs. In [NSV10] it was also proved that determinacy is undecidable if the elements of $\mathcal{Q}$ are CQs and Q_0 is a first order sentence (or the other way round). Another negative result is presented in [FGZ12]: determinacy is shown there to be undecidable if $\mathcal{Q}$ is a DATALOG program and Q_0 is CQ.

A. Our contribution. Finite vs. unrestricted case.

As we said, the case we study, where both the views and the query are CQs, is not the only one to be studied, but one of special importance, since CQs – as [NSV07] puts it – are "the simplest and most common language to define views and queries". The main technical result of this paper is:

Theorem 1: CQfDP is undecidable.

As usually in database theory one can consider two variants of the problem: *finite*, where all the structures in question (which in our case means $\mathbb{D}_1$ and $\mathbb{D}_2$) are assumed to be finite, and *unrestricted*, where there is no such assumption. Most of the results of [NSV07], [NSV10], [A11] and [P11] that we report above hold true regardless of the finiteness assumption.

A theorem analogous to Theorem 1 but concerning the unrestricted Conjunctive Query Determinacy Problem (CQDP) is the main result of our earlier paper [GM15]. Since the unrestricted case is viewed by the database theory community as less natural, it is fair to say that the result from [GM15] was perceived by the community as a step forward, but not as one closing the problem (see e.g. [F15]).

Since problems tend to be computationally harder in the finite case than in unrestricted, it was natural to conjecture, after [GM15], that Theorem 1 should hold true. But its proof is significantly more difficult than the respective proof in [GM15]. Let us try to explain why it is so.

As the reader is going to see in Section IV, determinacy (both finite and unrestricted) boils down to the question,

[1]This research was supported by Polish National Science Centre grant 2013/09/N/ST6/01188 (Tomasz Gogacz) and by Polish National Science Centre grant DEC-2013/09/B/ST6/01535 (Jerzy Marcinkowski).

[2]Actually, [F15] does such a good job as a survey that we will stick to an absolutely minimal introduction here, especially that we have enough of new technical material to easily overrun the page limit anyway.

whether some query (call it $red(Q_0)$) is true in **all** structures $\mathbb{M}$ (for the unrestricted case), or in all **finite** structures $\mathbb{M}$ (for the finite case) satisfying $\mathbb{M} \models \mathcal{T}_\mathcal{Q}$ and $\mathbb{M} \models green(Q_0)$, where $green(Q_0)$ is a structure, depending on Q_0, and $\mathcal{T}_\mathcal{Q}$ is a set of tuple generating dependencies, depending on $\mathcal{Q}$.

But in order to decide the above question for **all** $\mathbb{M}$ it is enough to study just **one** *universal structure*, namely $chase(\mathcal{T}_\mathcal{Q}, green(Q_0))$. Determinacy holds if and only if $red(Q_0)$ is true in this single structure.

This means that, when we encode some undecidable problem to show undecidability of (the unrestricted) CQDP, we only need to prove that whenever we start from a positive instance of our problem we get $\mathcal{Q}$ and Q_0 such that $chase(\mathcal{T}_\mathcal{Q}, green(Q_0)) \models red(Q_0)$, and whenever we start from a negative instance we get $\mathcal{Q}$ and Q_0 such that $chase(\mathcal{T}_\mathcal{Q}, green(Q_0)) \not\models red(Q_0)$

No such universal structure exists for the finite case. This generates problems of two sorts:

- How can we be sure (when we start from a negative instance[3]) that $\mathbb{M} \models red(Q_0)$ is true in all relevant finite structures $\mathbb{M}$? Do we have any technique that is specific for finite models?
- In order to show (when we start from a positive instance) that $\mathbb{M} \not\models red(Q_0)$, for some relevant $\mathbb{M}$, we need to built a finite model, for a set of TGDs, which omits some conjunctive query. We are dangerously close here to the issues related to finite controllability (see [R06], [BGO10], [GM13]) which are known to be difficult.

B. Our contribution. First Order non-rewritability

Let, as always, $\mathcal{Q} = \{Q_1 \dots Q_n\}$ be a set of CQs and let Q_0 be a CQ. Then, by (slightly restated) definition, $\mathcal{Q}$ (finitely) determines Q_0 if there is a function $h^{Q_0}_\mathcal{Q}$ which, for a (finite) structure $\mathbb{D}$ over Σ, takes, as its argument, the structure $\mathcal{Q}(\mathbb{D})$ and returns "yes" or "no", depending on whether $\mathbb{D} \models Q_0$ or not. Notice that $\mathcal{Q}(\mathbb{D})$ is no longer a structure over Σ. Its signature consists of one k-ary relation symbol for each query $Q_i \in \mathcal{Q}$ having k free variables.

It is known from [NSV07] that if $\mathcal{Q}$ determines Q_0 in the unrestricted sense then function $h^{Q_0}_\mathcal{Q}$ can be defined by a first order formula $\Theta^{Q_0}_\mathcal{Q}$, called *FO-rewriting* of Q_0 with respect to $\mathcal{Q}$: $h^{Q_0}_\mathcal{Q}(\mathcal{Q}(\mathbb{D})) = yes$ if and only if $\mathcal{Q}(\mathbb{D}) \models \Theta^{Q_0}_\mathcal{Q}$. The proof is via Craig Lemma, and since this lemma does not hold for First Order Logic over finite structures, it was natural to conjecture that the result will not survive if we restrict our attention to finite database instances. And indeed, as we are going to show:

Theorem 2: There are $\mathcal{Q}$ and Q_0 such that $\mathcal{Q}$ finitely determines Q_0 but $h^{Q_0}_\mathcal{Q}$ is not first order definable.

A detailed proof of this theorem is too long for a conference paper, and will only be presented in the full version of this paper. We however outline it here, and all the important proof ideas are already present in this outline.

[3]Notice that the roles of positive and negative instances have swapped. This is due to the fact that finite determinacy is co-r.e. and unrestricted determinacy is r.e.

II. Preliminaries

We need nothing beyond some standard finite model theory/database theory notions. They are only recalled here in order to fix notations.

A. Basic notions

When we say "structure" we mean a relational structure $\mathbb{D}$ over some signature Σ, i.e. a set of elements (vertices), denoted as $Dom(\mathbb{D})$ (or just as $\mathbb{D}$ if no confusion is possible) and a set of relational atoms, whose arguments are elements of $\mathbb{D}$ and whose predicate names are from Σ. Atoms are (of course) only positive. For an atomic formula A (or for any other formula) we use notation $\mathbb{D} \models A$ to say that A is true in $\mathbb{D}$.

Apart from predicate symbols Σ can also contain constants. If c is a constant from Σ and $\mathbb{D}$ is a structure over Σ then $c \in Dom(\mathbb{D})$.

$\mathbb{D}_1$ is a substructure of $\mathbb{D}$ (and $\mathbb{D}$ is a superstructure of $\mathbb{D}_1$) if for each atom A if $\mathbb{D}_1 \models A$ then $\mathbb{D} \models A$. This implies that $Dom(\mathbb{D}_1) \subseteq Dom(\mathbb{D})$.

For two structures $\mathbb{D}_1$ and $\mathbb{D}$ over the same signature Σ a function $h : Dom(\mathbb{D}_1) \to Dom(\mathbb{D})$ is called a homomorphism if for each $P \in \Sigma$ of arity l and each tuple $\bar{a} \in Dom(\mathbb{D}_1)^l$ if $\mathbb{D}_1 \models P(\bar{a})$ then $\mathbb{D} \models P(h(\bar{a}))$ (where $h(\bar{a})$ is a tuple of images of elements of $\bar{a}$).

A conjunctive query (over Σ), in short CQ, is a conjunction of atomic formulas (over Σ) whose arguments are either variables or the constants from Σ, preceded by existential quantifier binding some of the variables. It is important to distinguish between a CQ and its quantifier-free part.

For a conjunction of atoms Ψ (or for a CQ $Q(\bar{x}) = \exists \bar{y}\ \Psi(\bar{y}, \bar{x})$) the canonical structure of Ψ, denoted as $A[\Psi]$, is the structure whose elements are all the variables and constants appearing in Ψ and whose atoms are atoms of Ψ. It is useful to notice that for a finite structure $\mathbb{D}$ and a set $V \subseteq Dom(\mathbb{D})$ there is a unique conjunctive query Q such that $\mathbb{D} = A[Q]$ and that V is the set of free variables of Q.

For a CQ $Q(\bar{x}) = \exists \bar{y}\ \Psi(\bar{y}, \bar{x})$ with $\bar{x} = x_1, \dots x_l$, for a structure $\mathbb{D}$ and for a tuple $a_1, \dots a_l$ of elements of $\mathbb{D}$ we write $\mathbb{D} \models Q(a_1, \dots a_l)$ when there exists a homomorphism $h : A[\Psi] \to \mathbb{D}$ such that $h(x_i) = a_i$ for each i.

Sometimes we also write $\mathbb{D} \models Q$. Then we assume that all the free variables of Q are implicitly existentially quantified, so that the meaning of the notation is that there exists some homomorphism $h : A[\Psi] \to \mathbb{D}$.

The **most fundamental definition** of this paper now, needed to formulate the problem we solve: for a CQ Q and for a structure $\mathbb{D}$ by $Q(\mathbb{D})$ we denote the "view defined by Q over $\mathbb{D}$", which is the relation $\{\bar{a} : \mathbb{D} \models Q(\bar{a})\}$.

B. TGDs and how they act on a structure

A Tuple Generating Dependency (or TGD) is a formula of the form: $\forall \bar{x}, \bar{y}\ [\Phi(\bar{x}, \bar{y}) \Rightarrow \exists \bar{z}\ \Psi(\bar{z}, \bar{y})]$

where Ψ and Φ are conjunctions of atomic formulas. The standard convention, which we usually obey, is that the universal quantifiers in front of the TGD are omitted.

From the point of view of this paper it is important to see a TGD – let it be T, equal to $\Phi(\bar{x},\bar{y}) \Rightarrow \exists \bar{z}\ \Psi(\bar{z},\bar{y})$ – as a procedure whose input is a structure $\mathcal{D}$ and whose output is a new structure being a superstructure of $\mathcal{D}$:

find a tuple $\bar{b}$ (with $|\bar{b}| = |\bar{y}|$) such that:
① $\mathcal{D} \models \exists \bar{x}\ \Phi(\bar{x},\bar{b})$ via homomorphism h but
② $\mathcal{D} \not\models \exists z\ \Psi(\bar{z},\bar{b})$;

create a new copy of $A[\Psi]$;

output $\mathcal{D}(T,\bar{b})$ being a union of $\mathcal{D}$ and the new copy of $A[\Psi]$, with each y from $A[\Psi]$ identified with $h(y)$ in $\mathcal{D}$.

The message, which will be **good to remember**, is that the interface between the "new" part of the structure, added by a single application of a TGD to a structure, and the "old" structure, are the free variables of the query in the right hand side of the TGD.

C. Chase

For a set $\mathcal{T}$ of TGDs and for a structure $\mathbb{D}$ let $chase_0(\mathcal{T},\mathbb{D}) = \mathbb{D}$. Then, $chase_{i+1}(\mathcal{T},\mathbb{D})$ is defined by the procedure:

$\mathcal{D} := chase_i(\mathcal{T},\mathbb{D})$
forall pairs T, $\bar{b}$, where T is a TGD in $\mathcal{T}$ and $\bar{b}$ is a tuple of elements of $chase_i(\mathcal{T},\mathbb{D})$ **do:**
{ **if** ① and ② hold in $\mathcal{D}$ for $\bar{b}$ and T **then** $\mathcal{D} := \mathcal{D}(T,\bar{b})$ };
$chase_{i+1}(\mathcal{T},\mathbb{D}) := \mathcal{D}$.

Then $chase(\mathcal{T},\mathbb{D})$ is defined as $\bigcup_{i\in\mathbb{N}} chase_i(\mathcal{T},\mathbb{D})$. Notice that our chase is "lazy" – we only produce new atoms and new elements when needed.

III. Outline of the Technical Part

Most of this paper is devoted to the proof of Theorem 1.

Proving undecidability means encoding. Encoding means programming. Programming means the device that is being programmed, and the programming language.

In [GM15] we developed the device. An elementary "hardware" object there is a structure called *spider* and an elementary "instruction" able to act on a structure built out of spiders is a CQ called *spider query*.

In the Preliminaries we said that TGDs can "act on a structure". But how can a (conjunctive) query possibly do? This is explained in Section IV.

In Section V we try – without diving into details – to define an interface between the device from [GM15] and the current paper. In Section VI we define a high level programming language to manipulate spiders. We also define what is means to compile a program in such a language and show that this compilation is correct. This is new – the problem we encoded in [GM15] was much simpler than what we are going to encode here, and – using our running metaphor – we could afford programming in a language which was pretty low level. Some ideas of the proof of Lemma 12 were however present already in [GM15].

Separating example As we already have noticed, finite determinacy is co-r.e. and unrestricted determinacy is r.e. which implies – since we know from [GM15] that unrestricted determinacy is undecidable – that the two notions do not coincide. But no separating example was know so far. In Section VII we construct such an example. This is not just for curiosity. As it appears, this example is (together with Rainworm Machines) one of the two main engines of the proof of Theorem 1 which is presented in Section VIII.

Finally, in Section IX we explain the main ideas of the proof of Theorem 2. The separating example from Section VII turns out to also be a counterexample for FO-rewritability.

IV. Green-Red TGDs

A. Green-Red Signature

For a given signature Σ let Σ_G and Σ_R be two copies of Σ with new relation symbols, which have the same names and the same arities as symbols in Σ but are written in green and red respectively. Let $\bar{\Sigma}$ be the union of Σ_G and Σ_R. Notice that the constants from Σ (if there are any) are not relation symbols, so they are never colored and thus survive in $\bar{\Sigma}$ unharmed.

For any formula Ψ over Σ let $R(\Psi)$ (or $G(\Psi)$) be the result of painting all the predicates in Ψ red (green). For any formula Ψ over $\bar{\Sigma}$ let $dalt(\Psi)$ ("daltonisation of Ψ") be a formula over Σ being the result of erasing the colors from predicates of Ψ. The same convention applies to structures. For a structure $\mathcal{D}$ over $\bar{\Sigma}$, by $\mathcal{D}\restriction G$ we mean the substructure of $\mathcal{D}$ consisting of all its atoms which are over Σ_G. Analogously for $\mathcal{D}\restriction R$.

B. Having $\mathbb{D}$ instead of $\mathbb{D}_1$ and $\mathbb{D}_2$.

We restate CQfDP a little bit, as we prefer to be talking about one two-colored database instead of two. Clearly CQfDP can be equivalently restated as **CQfDP.2**:

> The instance of the problem is a finite set $\mathcal{Q}$ of conjunctive queries and another conjunctive query Q_0, all of them over some signature Σ. The question is whether for each finite structure $\mathbb{D}$ over $\bar{\Sigma}$ such that:
> ❶ $(G(Q))(\mathbb{D}) = (R(Q))(\mathbb{D})$ for each $Q \in \mathcal{Q}$
> it also holds that $(G(Q_0))(\mathbb{D}) = (R(Q_0))(\mathbb{D})$.

Definition 3: For a conjunctive query Q of the form $\exists \bar{x}\ \Phi(\bar{x},\bar{y})$ where Φ is a conjunction of atoms over Σ let $Q^{G\to R}$ be the TGD generated by Q *in the following sense:*

$$Q^{G\to R} \quad = \quad \forall \bar{x},\bar{y}\ [\ G(\Phi)(\bar{x},\bar{y}) \Rightarrow \exists \bar{z}\ R(\Phi)(\bar{z},\bar{y})\ ]$$

TGD $Q^{R\to G}$ is defined in an analogous way. For a set $\mathcal{Q}$ as above let $\mathcal{T}_\mathcal{Q}$ be the set of all TGDs of the form $Q^{G\to R}$ or $Q^{R\to G}$ with $Q \in \mathcal{Q}$. It is very easy to see that:

Lemma 4: The above condition ❶ is satisfied by structure $\mathbb{D}$ if and only if $\mathbb{D} \models \mathcal{T}_\mathcal{Q}$.

Now CQfDP.2 can be again restated as **CQfDP.3**:

Given a set $\mathcal{Q}$ (as in the formulation of CQfDP.2 above), and another conjunctive query Q_0, is it true that:

❷ for each finite structure $\mathbb{D}$ and each tuple $\bar{a}$ of elements of $\mathbb{D}$, if $\mathbb{D} \models \mathcal{T}_\mathcal{Q}, G(Q_0)(\bar{a})$ then also $\mathbb{D} \models R(Q_0)(\bar{a})$?

Of course the unrestricted version of CQfDP (called CQDP, see the Introduction) is equivalent to CQfDP.3 after removing, from its formulation, the word "finite".

The equivalent version of Theorem 1 which we actually prove in Sections VII and VIII is:

Theorem 5: **CQfDP.3** is undecidable.

V. Building on top of [GM15]: spiders and spider queries.

A. How to read this paper

This paper builds on top of the techniques developed in [GM15]. But we are of course not able to include here a presentation of the techniques from [GM15] which would be detailed enough to make the current paper self-contained. There are two (or three) possible ways of reading this paper:

- A good way is to first read Sections IV, V and VI.A of [GM15] (which is about 4 pages) and then jump to Section VI of the current paper.
- The shortest way is to read the next subsection, where we try to outline the ideas from [GM15] without going into details. We believe that most of the constructions of this paper should be understandable then, with the exception of the proof of Lemma 12(1), which constitutes an interface between the techniques from [GM15] and the new material.
- It is of course not at all forbidden to read both the next subsection and Sections IV, V and VI.A of [GM15] in any chosen order.

B. Spiders and spider queries – a crash course

As we observed in Section IV, conjunctive query (finite) determinacy is about (finite) structures being models of $\mathcal{T}_\mathcal{Q}$, for some set $\mathcal{Q}$ of CQs. So negative results concerning determinacy (in its both versions) are about writing nontrivial logic programs in the language having $\mathcal{T}_\mathcal{Q}$ as its set of instructions. $\mathcal{T}_\mathcal{Q}$ are TGDs, and normally encoding complicated things using TGDs should not be that difficult. But the TGDs from $\mathcal{T}_\mathcal{Q}$ are of a very special form, with right-hand side being merely an opposite-color version of the left-hand side. To understand to what extent this is a restriction it is good to realize that:

Observation 6 (Very easy): Let $\mathbb{D}$ be a structure over Σ_G and let $\mathcal{Q}$ be a set of CQs. Then there exists a homomorphism $h : dalt(chase(\mathcal{T}_\mathcal{Q}, \mathbb{D})) \to dalt(\mathbb{D})$.

This Observation seems to imply that $\mathcal{T}_\mathcal{Q}$ cannot produce anything "non-trivial" – we never get, in $dalt(chase(\mathcal{T}_\mathcal{Q}, \mathbb{D}))$, anything we did not have in $\mathbb{D}$ anyway[4].

[4]When we observed this phenomenon we were pretty sure we had just discovered a key fact in a decidability proof .

But, as we have discovered in [GM15], while daltonisation of $chase(\mathcal{T}_\mathcal{Q}, \mathbb{D})$ is indeed always an uncomplicated structure, there is a lot we can gain by playing with colors.

Let s be a natural number, large enough, and let $\mathbb{S} = \{1, 2, \dots s\}$. We defined a structure over Σ, which we call *spider*[5] and colored versions of spider: $\star$ and $\bigstar$ (green and red spiders). Each spider has $2s$ legs, including s "upper" and s "lower" legs, and $\star^I_J$ is a green spider with his I-th upper leg and J-th lower leg being red (analogously for $\bigstar^I_J$). $I, J \subseteq \mathbb{S}$ **are always** either singletons or empty, so we can have green spiders with none, one or two red legs[6] (and *vice versa*). The set of all spiders[7] of the form $\star^I_J$ or $\bigstar^I_J$ is denoted as $\mathbb{A}$.

Then we have the set $\mathbb{F}$ of **spider queries**. Elements of $\mathbb{F}$ are denoted as $\mathbb{f}^I_J$, where I and J are as always. The crucial observation is that the left hand side of the TGD $(\mathbb{f}^I_J)^{R\to G}$ matches with $\bigstar^{I'}_{J'}$ if and only if $I' \subseteq I$ and $J' \subseteq J$, and what is produced[8] by such an application of the TGD is $\star^{I\setminus I'}_{J\setminus J'}$. The same of course holds for the colors reversed. We like this trick so much that we call it the Rule of Spider Algebra:

$$\mathbb{f}^I_J(\bigstar^{I'}_{J'}) = \star^{I\setminus I'}_{J\setminus J'} \qquad (\clubsuit)$$

Binary queries. One more feature of a spider is that it has two vertices not involved in the mechanism enforcing the rule ♣. We call them *antenna* and *tail*. Also each query $\mathbb{f}^I_J$ has its antenna and its tail. Their role is to form – as we call them – *binary queries*.

The set $\mathbb{F}^2$ of binary queries contains, for each two queries $\mathbb{f}^I_J$ and $\mathbb{f}^{I'}_{J'}$ two conjunctive queries $\mathbb{f}^I_J \curlywedge \mathbb{f}^{I'}_{J'}$ and $\mathbb{f}^I_J \curlyvee \mathbb{f}^{I'}_{J'}$.

To define a conjunctive query one needs to specify its canonical structure and its set of free variables (which is a subset of the set of vertices of the canonical structure). The canonical structure of $\mathbb{f}^I_J \curlywedge \mathbb{f}^{I'}_{J'}$ is the disjoint union of canonical structures of $\mathbb{f}^I_J$ and of $\mathbb{f}^{I'}_{J'}$, with the only exception that the antennas of $\mathbb{f}^I_J$ and of $\mathbb{f}^{I'}_{J'}$ are identified. Their joint antennas are/is an existentially quantified variable in $\mathbb{f}^I_J \curlywedge \mathbb{f}^{I'}_{J'}$, and their tails are free variables in $\mathbb{f}^I_J \curlywedge \mathbb{f}^{I'}_{J'}$. All the remaining variables are free if and only if they were free in $\mathbb{f}^I_J$ or $\mathbb{f}^{I'}_{J'}$ – they do the magic of ♣ that we do not want to go into.

Now please come back to Definition 3 and notice that, when a query $Q = \mathbb{f}^I_J \curlywedge \mathbb{f}^{I'}_{J'}$ is seen as a green-red TGD $Q^{G\to R}$, the free variables in Q are what connects the new part of the structure, added by a single execution of $Q^{G\to R}$, to the old part. So, what a single execution of $Q^{G\to R}$ does is as follows: it finds, in the current structure, two green (real) spiders (call them $\mathcal{S}$ and $\mathcal{S}'$) with tails a and a', which share their antennas, and such that, according to ♣, $\mathbb{f}^I_J$ can be applied to $\mathcal{S}$ and $\mathbb{f}^{I'}_{J'}$

[5] Spider is parametrized by $s \in \mathbb{N}$, so in [GM15] we call it s-pider.

[6] $\star$ is short for $\star^\emptyset_\emptyset$.

[7]We should distinguish here between $2 + 4s + 2s^2$ "ideal spiders", which are elements of $\mathbb{A}$, and "real spiders" – homomorphic copies of the elements of $\mathbb{A}$ in some bigger structure.

[8]"Matches" means in particular, that the canonical structure of the query $\mathbb{f}^I_J$ is a substructure of the respective spider. "Produced" means that $\star^{I\setminus I'}_{J\setminus J'}$ emerges somehow in the structure, after the new vertices and new edges are added, as demanded by the right-hand side of the TGD $(\mathbb{f}^I_J)^{R\to G}$. Notice that all the new edges are green, so if there are any red edges in the resulting $\star^{I\setminus I'}_{J\setminus J'}$ they are inherited from the old structure.

can be applied to $\mathcal{S}'$. Then it creates two new (real) red spiders $\mathbb{f}^I_J(\mathcal{S})$ and $\mathbb{f}^{I'}_{J'}(\mathcal{S}')$, which again share the antennas, and their shared antenna is a new vertex. They are connected to the old structure via a, which is also the tail of the new $\mathbb{f}^I_J(\mathcal{S})$, via a' which is also the tail of the new $\mathbb{f}^{I'}_{J'}(\mathcal{S}')$ and via some vertices of the old structure which correspond to other (than the two tails) free variables of $\mathbb{f}^I_J \curlywedge \mathbb{f}^{I'}_{J'}$.

What concerns $\mathbb{f}^I_J \curlyvee \mathbb{f}^{I'}_{J'}$, again its canonical structure is the disjoint union of canonical structures of $\mathbb{f}^I_J$ and of $\mathbb{f}^{I'}_{J'}$. The only difference is that now the tails of $\mathbb{f}^I_J$ and of $\mathbb{f}^{I'}_{J'}$ are identified as one variable, and this variable is existentially quantified in $\mathbb{f}^I_J \curlyvee \mathbb{f}^{I'}_{J'}$. The two antennas are now free variables. The way $(\mathbb{f}^I_J \curlyvee \mathbb{f}^{I'}_{J'})^{G\to R}$ acts on a structure is analogous to the one of $(\mathbb{f}^I_J \curlywedge \mathbb{f}^{I'}_{J'})^{G\to R}$, which was explained above. Same for $(\mathbb{f}^I_J \curlywedge \mathbb{f}^{I'}_{J'})^{R\to G}$ and $(\mathbb{f}^I_J \curlyvee \mathbb{f}^{I'}_{J'})^{R\to G}$.

VI. Climbing the abstraction ladder

The language of spiders, whose signature is $\bar{\Sigma}$, which we briefly described in the previous section, is a very low level one – we think it is Abstraction Level Zero. But in order to show our undecidability result we need to produce pretty complicated programs in the language of spider queries from $\mathbb{F}^2$. The typical Computer Science way in such situation is to define a more abstract, higher order language, use it as the actual programming language (so that one does not need to worry about low level implementation details), and then compile the program written in this higher order language into the executable low-level form

Defining such a higher order language (or rather languages – we will first *precompile* the original program into an intermediate language and then *compile*) is precisely what we are going to do in this section. Each of the two languages we are going to define will comprise a relational signature and a set of graph rewriting rules (being TGDs, in disguise) which will act on structures over this signature.

Abstraction Level 1 language. The signature consists of one binary relation $H(\mathcal{S},\_,\_)$ for each (ideal) spider $\mathcal{S}\in\mathbb{A}$. A structure over this signature will be called a *swarm*. Now we are going to define the set $\mathbb{L}_1$ of *swarm rewriting rules*.

Definition 7: For each query $\mathbb{f}^{I_1}_{J_1} \curlyvee \mathbb{f}^{I_2}_{J_2}$ from $\mathbb{F}^2$ (and for each query $\mathbb{f}^{I_1}_{J_1} \curlywedge \mathbb{f}^{I_2}_{J_2}$ from $\mathbb{F}^2$) there will be a rule $\mathbb{f}^{I_1}_{J_1} \curlyvee \mathbb{f}^{I_2}_{J_2}$ (resp. $\mathbb{f}^{I_1}_{J_1} \curlywedge \mathbb{f}^{I_2}_{J_2}$) in $\mathbb{L}_1$, where $\mathbb{f}^{I_1}_{J_1} \curlyvee \mathbb{f}^{I_2}_{J_2}$ is a shorthand of:

$$\bigwedge_{I'_1\subseteq I_1, J'_1\subseteq J_1, I'_2\subseteq I_2, J'_2\subseteq J_2}$$
$$[\forall x,y,y'\ \ H(\bigstar^{I'_1}_{J'_1},x,y)\wedge H(\bigstar^{I'_2}_{J'_2},x,y')\Rightarrow \exists x'\ \ H(\bigstar^{I_1\setminus I'_1}_{J_1\setminus J'_1},x',y)\wedge H(\bigstar^{I_2\setminus I'_2}_{J_2\setminus J'_2},x',y')]$$
$$\wedge$$
$$[\forall x,y,y'\ \ H(\bigstar^{I'_1}_{J'_1},x,y)\wedge H(\bigstar^{I'_2}_{J'_2},x,y')\Rightarrow \exists x'\ \ H(\bigstar^{I_1\setminus I'_1}_{J_1\setminus J'_1},x',y)\wedge H(\bigstar^{I_2\setminus I'_2}_{J_2\setminus J'_2},x',y')]$$

and $\mathbb{f}^{I_1}_{J_1} \curlywedge \mathbb{f}^{I_2}_{J_2}$ is a shorthand of the formula:

$$\bigwedge_{I'_1\subseteq I_1, J'_1\subseteq J_1, I'_2\subseteq I_2, J'_2\subseteq J_2}$$
$$[\forall x,x',y\ \ H(\bigstar^{I'_1}_{J'_1},x,y)\wedge H(\bigstar^{I'_2}_{J'_2},x',y)]\Rightarrow \exists y'\ H(\bigstar^{I_1\setminus I'_1}_{J_1\setminus J'_1},x,y')\wedge H(\bigstar^{I_2\setminus I'_2}_{J_2\setminus J'_2},x',y')]$$
$$\wedge$$
$$[\forall x,x',y\ \ H(\bigstar^{I'_1}_{J'_1},x,y)\wedge H(\bigstar^{I'_2}_{J'_2},x',y)\Rightarrow \exists y'\ H(\bigstar^{I_1\setminus I'_1}_{J_1\setminus J'_1},x,y')\wedge H(\bigstar^{I_2\setminus I'_2}_{J_2\setminus J'_2},x',y')]$$

Horrible. But this is only because we wrote the rules as FOL formulas, while they are actually easy to explain in the natural language: $\mathbb{f}^{I_1}_{J_1} \curlyvee \mathbb{f}^{I_2}_{J_2}$ means that whenever two edges can be found in the current swarm, leading from x to y and from x to y', labelled with two spiders $\mathcal{S}_1$ and $\mathcal{S}_2$, of the same color, such that, according to the rule of Spider Algebra ♣, query $\mathbb{f}^{I_1}_{J_1}$ can be applied to $\mathcal{S}_1$ and $\mathbb{f}^{I_2}_{J_2}$ can be applied to $\mathcal{S}_2$, there must be also a vertex x' in this swarm, with edges from x' to y and from x' to y', labelled with spiders $\mathbb{f}^{I_1}_{J_1}(\mathcal{S}_1)$ and $\mathbb{f}^{I_2}_{J_2}(\mathcal{S}_2)$. And analogously for $\mathbb{f}^{I_1}_{J_1} \curlywedge \mathbb{f}^{I_2}_{J_2}$.

So, at Abstraction Level One we no longer need to think about the details of spider anatomy but we are still constrained by the, hardly intuitive, rule ♣. At Abstraction Level 2 we are going to liberate ourselves also from this constraint.

Abstraction Level 2 language. Let now $\mathbb{A}_2$ be the subset of $\mathbb{A}$ consisting of all green (ideal) spiders which are of the form $\bigstar^I$. The signature of Abstraction Level 2 language consists of one binary relation $H(\mathcal{S},\_,\_)$ for each $\mathcal{S}\in\mathbb{A}_2$. A structure over this signature will be called a *green graph*.

Notice that there is a natural bijection between $\mathbb{A}_2$ and $\bar{\mathbb{S}} = \mathbb{S}\cup\{\emptyset\}$, so we will often write[9] $H_i(x,y)$ instead of $H(\bigstar^{\{i\}},x,y)$ and $H_\emptyset(x,y)$ instead of $H(\bigstar,x,y)$.

Concerning the set $\mathbb{L}_2$, of *green graph rewriting rules*, for each four spiders $\bigstar^{I_1}\neq\bigstar^{I_3}$, $\bigstar^{I_2}\neq\bigstar^{I_4}\in\mathbb{A}_2$, there will be two rules in the set $\mathbb{L}_2$, denoted as: $I_1 \curlywedge I_2 \leftrightarrow I_3 \curlywedge I_4$ and as $I_1 \curlyvee I_2 \leftrightarrow I_3 \curlyvee I_4$ where $I_1 \curlywedge I_2 \leftrightarrow I_3 \curlywedge I_4$ is a shorthand of:

$$\forall x,x'\quad [\exists y\ \ H(\bigstar^{I_1},x,y)\wedge H(\bigstar^{I_2},x',y)]\Leftrightarrow [\exists y\ \ H(\bigstar^{I_3},x,y)\wedge H(\bigstar^{I_4},x',y)]$$

and $I_1 \curlyvee I_2 \leftrightarrow I_3 \curlyvee I_4$ is a shorthand of:

$$\forall y,y'\quad [\exists x\ \ H(\bigstar^{I_1},x,y)\wedge H(\bigstar^{I_2},x,y')]\Leftrightarrow [\exists x\ \ H(\bigstar^{I_3},x,y)\wedge H(\bigstar^{I_4},x,y')]$$

From now on it will be assumed that spiders $\bigstar^3$ and $\bigstar^4$ (that is – sets $\{3\}$ and $\{4\}$) do not occur in our sets of green graph rewriting rules.

Compilation and its correctness. Swarm rewriting rules, as well as green graph rewriting rules, are first order sentences, and each of them is equivalent to a conjunction of tuple generating dependencies. So, for a set $\mathcal{T}\subseteq\mathbb{L}_1$ (or $\mathcal{T}\subseteq\mathbb{L}_2$) and a swarm (resp. green graph) $\mathbb{D}$ the statement $\mathbb{D}\models\mathcal{T}$ makes sense and also the notion of Chase applies to $\mathcal{T}$.

Definition 8: For a set $\mathcal{T}\subseteq\mathbb{L}_1$ let Compile$(\mathcal{T})$ $=$ $= \{f \curlywedge f' : f \curlywedge f'\in\mathcal{T}\}\cup\{f \curlyvee f' : f \curlyvee f'\in\mathcal{T}\}$.

[9] Having two alternative notations for the same object looks like asking for confusion. But see: when discussing Precompilation we need to relate Level Two language to Level One, so it is natural to use the notation where spiders are explicit. But then, in Sections VII and VIII we do not need spiders any more, and we would hate to have the complicated tuples there as small subscripts. So $\bar{\mathbb{S}}$ fits us better than $\mathbb{A}_2$ there.

Which means "treat each rule from $\mathcal{T}$ as a binary query from $\mathbb{F}^2$".

Definition 9: For a set $\mathcal{T} \subseteq \mathbb{L}_2$ we define Precompile($\mathcal{T}$) $\subseteq \mathbb{L}_1$ as the result of the following procedure:

- *Precompile($\mathcal{T}$) $:= \{\mathbb{f}_1^1 \curlywedge \mathbb{f}_2^2, \mathbb{f}_1^3 \curlywedge \mathbb{f}_2^4, \mathbb{f}^3 \curlywedge \mathbb{f}_3^4\}$*
- *fix any numbering of the rules of $\mathcal{T}$ using natural numbers $2, 3 ... k$;*
- *for $i = 2$ to k,*

if the i'th rule in $\mathcal{T}$ is $I_1 \curlywedge I_2 \leftrightarrow I_3 \curlywedge I_4$ then add to Precompile($\mathcal{T}$) the rules $\mathbb{f}_{2i+1}^{I_1} \curlywedge \mathbb{f}_{2i+2}^{I_2}$ and $\mathbb{f}_{2i+1}^{I_3} \curlywedge \mathbb{f}_{2i+2}^{I_4}$
and if the i'th rule in $\mathcal{T}$ is $I_1 \curlyvee I_2 \leftrightarrow I_3 \curlyvee I_4$ then add to Precompile($\mathcal{T}$) the rules $\mathbb{f}_{2i+1}^{I_1} \curlyvee \mathbb{f}_{2i+2}^{I_2}$ and $\mathbb{f}_{2i+1}^{I_3} \curlyvee \mathbb{f}_{2i+2}^{I_4}$

Remark 10: The idea behind rules of the form $\mathbb{f}_{2i+1}^{I_1} \curlywedge \mathbb{f}_{2i+2}^{I_2}$ and $\mathbb{f}_{2i+1}^{I_3} \curlywedge \mathbb{f}_{2i+2}^{I_4}$ in Precompile($\mathcal{T}$) is that they simulate (in a swarm), in two steps, one execution of $I_1 \curlywedge I_2 \leftrightarrow I_3 \curlywedge I_4$ (in a green graph). As a by-product, two red edges are added to the swarm – labelled with $\bigstar_{2i+1}$ and $\bigstar_{2i+2}$ (the same for $\curlyvee$ instead of $\curlywedge$).

Definition 11:

- *Let $\mathcal{Q} \subseteq \mathbb{F}^2$. We will say that $\mathcal{Q}$* leads to the red spider *(or* finitely leads to the red spider*) if and only if each (resp. each finite) structure $\mathbb{D}$ over $\bar{\Sigma}$, such that $\mathbb{D} \models \mathcal{T}_\mathcal{Q}$, which contains a copy of the full green spider $\bigstar$, also contains a copy of the full red spider $\bigstar$.*
- *Let $\mathcal{T} \subseteq \mathbb{L}_1$. We will say that $\mathcal{T}$* leads to the red spider *(or* finitely leads to the red spider*) if each (resp. each finite) swarm $\mathbb{D}$ such that $\mathbb{D} \models \mathcal{T}$, which contains an atom of the relation $H(\bigstar, \_, \_)$, also contains an atom of the relation $H(\bigstar, \_, \_)$.*
- *We say that a green graph contains a 1-2 pattern if contains edges $H(\bigstar^1, a, b)$ and $H(\bigstar^2, a', b)$ for some vertices a, a', b.*
- *Let $\mathcal{T} \subseteq \mathbb{L}_2$. We will say that $\mathcal{T}$* leads to the red spider *(or* finitely leads to the red spider*) if each (resp. each finite) green graph $\mathbb{D}$ such that $\mathbb{D} \models \mathcal{T}$ which contains an atom of the relation $H(\bigstar, \_, \_)$ also contains a 1-2 pattern*[10].

Lemma 12:
1) *Let $\mathcal{T} \subseteq \mathbb{L}_1$. Then $\mathcal{T}$* leads to the red spider *(or finitely leads to the red spider) if and only if Compile($\mathcal{T}$) does.*
2) *Let $\mathcal{T} \subseteq \mathbb{L}_2$. Then $\mathcal{T}$* leads to the red spider *(or finitely leads to the red spider) if and only if Precompile($\mathcal{T}$) does.*

Proof of the Lemma can be found in Appendix A.

It is important to notice (see condition ❷ in Section IV) that:

Observation 13: A set of queries $\mathcal{Q} \subseteq \mathbb{F}^2$ leads (finitely leads) to the red spider if and only if it (finitely) determines $\exists^* dalt(\bigstar)$ (where $\exists^*$ means that all free variables are quantified, leading to a boolean query). So, by Lemma 12, since both Precompilation and Compilation are computable, in order to prove Theorem 5, it suffices to show that it is undecidable for a set $\mathcal{T} \subseteq \mathbb{L}_2$ whether $\mathcal{T}$ finitely leads to the red spider.

From now on proof of Theorem 5 has nothing to do with spiders. It is all about green graphs and their rewriting rules.

VII. A SEPARATING EXAMPLE

In this Section we are going to prove:

Theorem 14: There exists a set $\mathfrak{T} \subseteq \mathbb{L}_2$ of green graph rewriting rules which does not lead to the red spider, but finitely leads to the red spider.

Notice that, by Lemma 12, this will imply that the set $Compile(Precompile(\mathfrak{T}))$ of conjunctive queries over Σ does not determine the query $\exists^* dalt(\bigstar)$ but finitely determines it.

Here is how we are going to construct $\mathfrak{T}$:

Step 1. Let $\mathbb{D}_\bigstar$ be a green graph containing just two vertices $\mathbf{a}, \mathbf{b}$ and one edge $H_\emptyset(\mathbf{a}, \mathbf{b})$ (the constants $\mathbf{a}$ and $\mathbf{b}$ from $\mathbb{D}_\bigstar$ will be important, please befriend them). We will define a set $\mathfrak{T}_\infty$ of green graph rewriting rules such that $chase(\mathfrak{T}_\infty, \mathbb{D}_\bigstar)$ is an infinite green graph (a sort of *infinite path*) without a 1-2 pattern.

Step 2. Another set $\mathfrak{T}_\boxplus$ of green graph rewriting rules will be constructed, and $\mathfrak{T}$ will be defined as the union of $\mathfrak{T}_\infty$ and $\mathfrak{T}_\boxplus$. The rules of $\mathfrak{T}_\boxplus$ will be quite complicated (or at least numerous) and it will follow from the construction that $\mathfrak{T}$ finitely leads to the red spider.

Step 3. We will construct (an infinite) green graph $\mathbb{M}$, containing $\mathbb{D}_\bigstar$, without the 1-2 pattern, and such that $\mathbb{M} \models \mathfrak{T}$.

Step 1. Let $\mathfrak{T}_\infty$ consist of three green graph rewriting rules: (I) $\emptyset \curlywedge \emptyset \leftrightarrow \alpha \curlywedge \eta_1$ (II) $\emptyset \curlyvee \eta_1 \leftrightarrow \eta_0 \curlyvee \beta_1$ (III) $\emptyset \curlywedge \eta_0 \leftrightarrow \eta_1 \curlywedge \beta_0$ where α, β_0 and η_0 are some even numbers from $\mathbb{S}$, and β_1 and η_1 are odd.

Let us try to construct $chase(\mathfrak{T}_\infty, \mathbb{D}_\bigstar)$. We begin from a single edge $H_\emptyset(\mathbf{a}, \mathbf{b})$. The only rule that can be applied to this structure is (I). A new vertex b_1 is then created, together with two edges: $H_\alpha(\mathbf{a}, b_1)$ and $H_{\eta_1}(\mathbf{a}, b_1)$. In the next step the only way is to use rule (II), as we have $H_\emptyset(\mathbf{a}, \mathbf{b})$ and $H_{\eta_1}(\mathbf{a}, b_1)$ sharing the beginning vertex. We get a new vertex a_1 and new edges $H_{\eta_0}(a_1, \mathbf{b})$ and $H_{\beta_1}(a_1, b_1)$. In the third step we can apply rule (III) to $H_\emptyset(\mathbf{a}, \mathbf{b})$ and $H_{\eta_0}(a_1, \mathbf{b})$, creating b_2 with $H_{\eta_1}(\mathbf{a}, b_2)$ and $H_{\beta_0}(a_1, b_2)$. And so on – using rules (II) and (III) alternately we construct (see Fig. 1) infinite sequences b_1, $b_2 \ldots$ of vertices with out-degree 0 and a_1, $a_2 \ldots$ of vertices with in-degree 0, connected, in a regular way, with edges labelled with β_0 and β_1.

Seeing a green graph as a set of words (in parity glasses). It is standard to see paths in a directed labelled graph as words:

Definition 15: Let $\mathfrak{M}$ be a green graph with s, t among its vertices. Then paths($\mathfrak{M}, s, t$) is the set of all such words $\mathfrak{w} \in \mathbb{S}^*$ that, had $\mathfrak{M}$ been a nondeterministic finite automaton, with the initial state s and a single accepting state t, it would accept $\mathfrak{w}$, but it would not accept any nonempty proper prefix of $\mathfrak{w}$.

[10] Notice that it would make no sense to repeat, for $\mathcal{T} \subseteq \mathbb{L}_2$, the definition for $\mathcal{T} \subseteq \mathbb{L}_1$. This is because a green graph never has any atom of the relation $H(\bigstar, \_, \_)$. But (and please see it as an exercise) having a 1-2 pattern available, the set of rules $Precompile(\mathcal{T})$ will produce, in three steps, using rules $\mathbb{f}_1^1 \curlywedge \mathbb{f}_2^2$, $\mathbb{f}_1^3 \curlywedge \mathbb{f}_2^4$, $\mathbb{f}^3 \curlywedge \mathbb{f}_3^4$, an atom of $H(\bigstar, \_, \_)$.

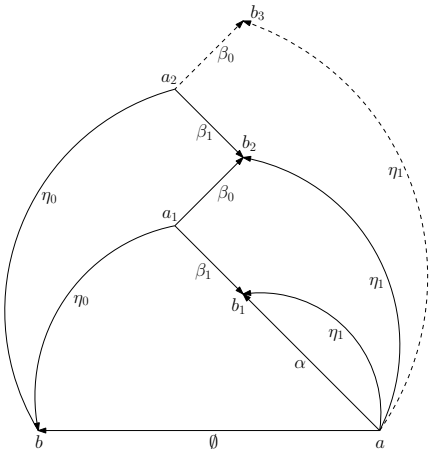

Figure 1. The structure $chase(\mathfrak{T}_\infty, \mathbb{D}_\star)$ *in statu nascendi*. Notice (this will be important in Section IX) that $chase_{i+1}(\mathfrak{T}_\infty, \mathbb{D}_\star)$ is always a result of exactly one application of a rule from $\mathfrak{T}_\infty$ to elements of $chase_i(\mathfrak{T}_\infty, \mathbb{D}_\star)$.

It will be crucial for us to see green graphs as complicated sets of words. But unfortunately in the interesting graphs, all directed paths will have length 1. Each vertex will either have out-degree 0 or in-degree 0 (see Figure 1). This is why we need Parity Glasses (reminder: elements of $\mathbb{S}$ are natural numbers).

Definition 16: Let $\mathfrak{M}$ be a green graph containing $\mathbb{D}_\star$ (which means that $\mathfrak{M}$ has an edge $H_\star(\mathbf{a}, \mathbf{b})$). Then:

- *PG($\mathfrak{M}$) is the graph resulting from $\mathfrak{M}$ by:*
  1) *Removing all the edges labelled with $\emptyset$;*
  2) *reversing the direction of all edges labelled with odd numbers.*
- $words(\mathfrak{M}) = paths(PG(\mathfrak{M}), \mathbf{a}, \mathbf{a}) \cup paths(PG(\mathfrak{M}), \mathbf{a}, \mathbf{b})$.

Example. $words(chase(\mathfrak{T}_\infty, \mathbb{D}_\star)) =$
$= \{\alpha(\beta_1\beta_0)^k\eta_1 : k \in \mathbb{N}\} \cup \{\alpha(\beta_1\beta_0)^k\beta_1\eta_0 : k \in \mathbb{N}\}$

By an $\alpha\beta$-path in a green graph $\mathbb{M}$ we mean a sequence of edges (or vertices, it will be always clear from the context), which seen as a word in $PG(\mathbb{M})$ is of the form $\alpha(\beta_1\beta_0)^*$. There are infinitely many $\alpha\beta$-paths in $chase(\mathfrak{T}_\infty, \mathbb{D}_\star)$, including for example the path $\mathbf{a}, b_1, a_1, b_2$ and the path $\mathbf{a}, b_1, a_1, b_2, a_3, b_3$ (see Figure 1).

Step 2. When a set $\mathcal{T}$ of green graph rewriting rules leads to the red spider then this fact is – at least in principle – easy to prove. One just builds $chase(\mathcal{T}, \mathbb{D}_\star)$ until a 1-2 pattern emerges. But how could we possibly prove that $\mathfrak{T}$ (which still remains to be defined, but which is **not** going to lead to the red spider) does **finitely** lead to the red spider?

The last means that a 1-2 pattern must emerge in *every* finite model of $\mathcal{T}$ containing $\mathbb{D}_\star$. Is there anything that we know for sure about *every* finite model of $\mathcal{T}$ containing $\mathbb{D}_\star$? Yes, we know[11] that for each such model $\mathbb{M}$ there exists a homomorphism $h : chase(\mathcal{T}, \mathbb{D}_\star) \to \mathbb{M}$. Our $\mathfrak{T}$ is going to be a superset of $\mathfrak{T}_\infty$ so $chase(\mathfrak{T}, \mathbb{D}_\star)$ will contain, as a substructure, the structure $chase(\mathfrak{T}_\infty, \mathbb{D}_\star)$ that we analyzed in Step 1. This means that there for sure will be two vertices $b_t \neq b_{t'}$, such that (remember – $\mathbb{M}$ is finite) $h(b_t) = h(b_{t'})$. This means that there are two $\alpha\beta$-paths in $\mathbb{M}$: $h(a), h(b_1), h(a_1) \ldots h(b_t)$ and $h(a), h(b_1), h(a_1) \ldots h(b_{t'})$, of **different lengths**, which share the endpoint (see Figure 2, do not look at the grid yet).

[11]The fact that Chase is a universal structure is one of the textbook facts of database theory [JK82].

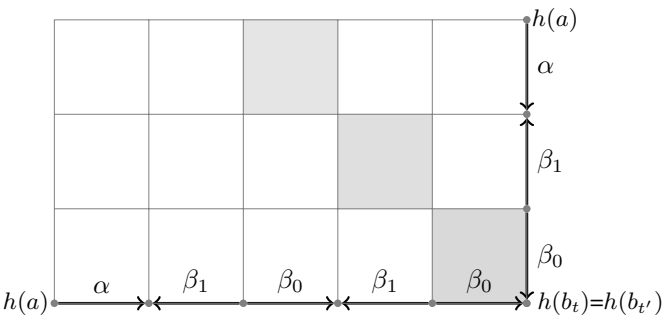

Figure 2. Long initial fragments of the two paths are also equal in $\mathbb{M}$ but we are never going to use this fact, and it is not indicated on the picture.

The set $\mathfrak{T}_\boxplus$ will be designed to detect such two $\alpha\beta$-paths and – after detecting them – to create a 1-2 pattern. Figure 2 explains how this will be done: the rules of $\mathfrak{T}_\boxplus$ will build a grid whose eastern border will be the $\alpha\beta$-path $h(\mathbf{a}), h(b_1), h(a_1) \ldots h(b_t)$ and whose southern border will be the path $h(\mathbf{a}), h(b_1), h(a_1) \ldots h(b_{t'})$.

This will be of course done step by step – as a result of a single application of a rule from $\mathfrak{T}_\boxplus$ one little square of the grid will be created. For such a step a rule will need the southern and eastern edge of the little square to exist, and it will add the western and northern edge[12]. After the complete grid is built the rules of $\mathfrak{T}$ will somehow check whether the northwestern corner of the grid is on the diagonal of the grid. If it is not, then indeed the $\alpha\beta$-paths $h(\mathbf{a}), h(b_1), h(a_1) \ldots h(b_t)$ and $h(\mathbf{a}), h(b_1), h(a_1) \ldots h(b_{t'})$ were of different lengths.

While the idea is simple, for the real construction we need 41 green graph rewriting rules and $4 \times 2^3 = 32$ binary relations for the inner edges of the grid (by which we mean edges not belonging to one of the two $\alpha\beta$-paths, so that inner vertices, in our sense, also include western and northern border). The names of the 32 relations (or labels of the green graphs edges) will be conveniently encoded[13] as $\langle n|e|s|w, \alpha|\beta, d|\bar{d}, b|\bar{b}\rangle$ (where | is the BNF "or"). We think that $\langle n, \alpha, \bar{d}, \bar{b}\rangle$ is 1 and $\langle w, \alpha, \bar{d}, \bar{b}\rangle$ is 2, where 1 and 2 are the ones from 1-2-pattern.

The first parameter of a label of en edge – one of n, e, s, w – is the direction the edge heads. The second – α or β – is inherited from the "respective" element of one of the original $\alpha\beta$-paths. The parameter d (or $\bar{d}$) tells us whether one of the ends of the edge is (or is not) on the diagonal of the grid (see Figure 3). The fourth parameter (needed in Step 3) tells if an

[12]See – adding two missing edges of a square is exactly what green graph rewriting rules are good at.

[13] Which, precisely speaking, means that we assume there is some fixed bijection between our new set of codes and some subset of $\mathbb{S}$ and that we identify a code with its image under this bijection. We of course assume that this subset is disjoint with $\{\alpha, \beta_0, \beta_1, \eta_0, \eta_1\}$.

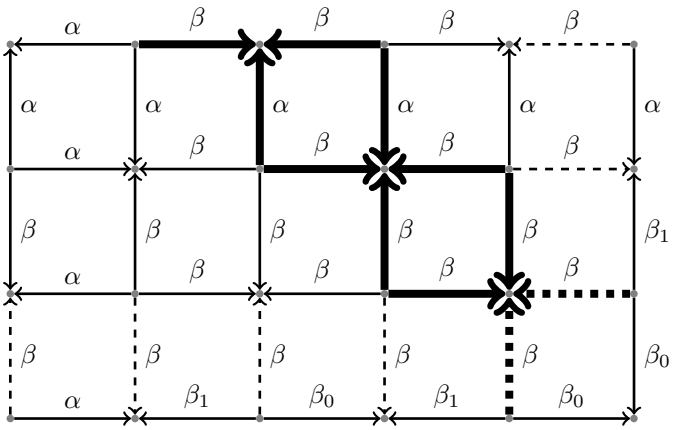

Figure 3. Grid constructed by $\mathfrak{T}_{\boxplus}$. Diagonal ("d") edges are bold and border ("b") edges are dashed. Labels of the edges in the NW-corner are $\langle n, \alpha, \bar{d}, \bar{b}\rangle$ and $\langle w, \alpha, \bar{d}, \bar{b}\rangle$ so they form a 1-2 pattern.

edge shares a vertex with one of the original $\alpha\beta$-paths.

Now, once we understand the sense of the parameters, it is time to see **the rules of** $\mathfrak{T}_{\boxplus}$:

$\beta_0 \fatbslash \beta_0 \leftrightarrow \langle n, \beta, d, b\rangle \fatbslash \langle w, \beta, d, b\rangle$

The above rule (call it *grid triggering rule*) creates the tile in the south-eastern corner of the grid. The next four are:

$\beta_1 \fatslash \langle n, \beta, d, b\rangle \leftrightarrow \langle s, \beta, \bar{d}, b\rangle \fatslash \langle e, \beta, d, \bar{b}\rangle$
$\beta_0 \fatbslash \langle s, \beta, \bar{d}, b\rangle \leftrightarrow \langle n, \beta, \bar{d}, b\rangle \fatbslash \langle w, \beta, \bar{d}, \bar{b}\rangle$
$\beta_1 \fatslash \langle n, \beta, \bar{d}, b\rangle \leftrightarrow \langle s, \beta, \bar{d}, b\rangle \fatslash \langle e, \beta, \bar{d}, \bar{b}\rangle$
$\alpha \fatbslash \langle s, \beta, \bar{d}, b\rangle \leftrightarrow \langle n, \beta, \bar{d}, b\rangle \fatbslash \langle w, \alpha, \bar{d}, \bar{b}\rangle$

They build the strip of tiles adjacent to the southern edge of the rectangle. Analogously, the strip of tiles adjacent to the eastern edge of the rectangle will be built by the four rules:

$\beta_1 \fatslash \langle w, \beta, d, b\rangle \leftrightarrow \langle e, \beta, \bar{d}, b\rangle \fatslash \langle s, \beta, d, \bar{b}\rangle$
$\beta_0 \fatbslash \langle e, \beta, \bar{d}, b\rangle \leftrightarrow \langle w, \beta, \bar{d}, b\rangle \fatbslash \langle n, \beta, \bar{d}, \bar{b}\rangle$
$\beta_1 \fatslash \langle w, \beta, \bar{d}, b\rangle \leftrightarrow \langle e, \beta, \bar{d}, b\rangle \fatslash \langle s, \beta, \bar{d}, \bar{b}\rangle$
$\alpha \fatbslash \langle w, \beta, \bar{d}, b\rangle \leftrightarrow \langle w, \beta, \bar{d}, b\rangle \fatbslash \langle n, \alpha, \bar{d}, \bar{b}\rangle$

Now last 32 rules of $\mathfrak{T}_{\boxplus}$ are coming, which will build the interior of the rectangle (including the strips adjacent to the northern and western borders). Due to the space limit we write them as two schemes, each representing 16 rules:

$\langle e, \Theta, X, \bar{b}\rangle \fatbslash \langle s, \Omega, Y, \bar{b}\rangle \leftrightarrow \langle n, \Omega, X, \bar{b}\rangle \fatbslash \langle w, \Theta, Y, \bar{b}\rangle$
for each $X, Y \in \{d, \bar{d}\}$ and for each $\Theta, \Omega \in \{\alpha, \beta\}$.

$\langle w, \Theta, X, \bar{b}\rangle \fatslash \langle n, \Omega, Y, \bar{b}\rangle \leftrightarrow \langle s, \Omega, X, \bar{b}\rangle \fatslash \langle e, \Theta, Y, \bar{b}\rangle$
for each $X, Y \in \{d, \bar{d}\}$ and for each $\Theta, \Omega \in \{\alpha, \beta\}$.

Now it follows from the construction that:

Lemma 17: $\mathfrak{T} = \mathfrak{T}_{\boxplus} \cup \mathfrak{T}_{\infty}$ *finitely leads to the red spider.*

Step 3. Since now we are going to build an infinite model $\mathbb{M}$ for $\mathfrak{T}$, we can afford having $chase(\mathfrak{T}_{\infty}, \mathbb{D}_{\fatstar})$ as a substructure of $\mathbb{M}$ and we do not need to identify any of its vertices. But (unfortunately) the grid triggering rule still can be applied – this is since the edges labels in the left hand side of the first rule of $\mathfrak{T}_{\boxplus}$ are equal, which means that for each edge $H_{\beta_0}(a_t, b_{t+1})$ in $chase(\mathfrak{T}_{\infty}, \mathbb{D}_{\fatstar})$ a new vertex, call it c_t and two new edges, both leading from a_t to c_t, will be created, namely $H_{\langle n,\beta,d,b\rangle}(a_t, c_t)$ and $H_{\langle w,\beta,d,b\rangle}(a_t, c_t)$. Then the construction from **Step 2** can be repeated, leading to a new

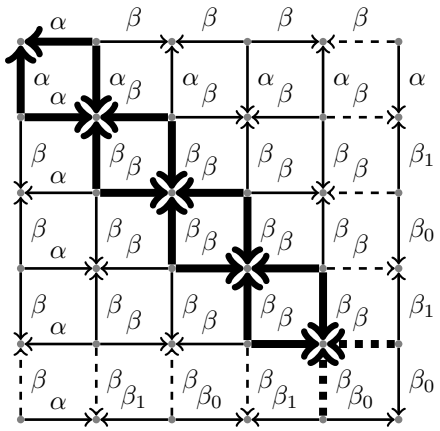

Figure 4. Grid $\mathbb{M}_3$. No 1-2 pattern.

grid $\mathbb{M}_t$ (see Figure 4), constructed in the way described in **Step 2** but without a 1-2 pattern. Notice that the picture does not fully reflect the reality here[14] – in the real green graph the respective vertices of the southern and eastern borders of $\mathbb{M}_t$ are equal.

Each of $\mathbb{M}_t$ contains some vertices and atoms of $chase(\mathfrak{T}_{\infty}, \mathbb{D}_{\fatstar})$ so, for $t \leq t'$ the intersection of $\mathbb{M}_t$ and $\mathbb{M}_{t'}$ is exactly the $\alpha\beta$-path $\mathbf{a}, b_1, a_1 \dots b_t$. Define $\mathbb{M} = chase(\mathfrak{T}_{\infty}, \mathbb{D}_{\fatstar}) \cup \bigcup_{t\in\mathbb{N}} \mathbb{M}_t$. Theorem 14 follows from:

Lemma 18: 1) $\mathbb{M}$ *does not contain a 1-2 pattern.*
2) $\mathbb{M} \models \mathfrak{T}$

For the (easy) proof of Lemma 18 see Appendix B.

VIII. Proof of Theorem 5

In $\mathfrak{T}_{\infty}$ we had η_0 and η_1 calling each other in an infinite loop, and creating an infinite $\alpha\beta$-path. By homomorphism argument, in presence of such path, the rules of $\mathfrak{T}_{\boxplus}$ lead to a 1-2 pattern. Now this simple mutual recursive call will be controlled by something undecidably complicated – a rainworm.

A. Rainworm machine. And how it creeps.

Rainworm machine (RM), which we will now define, is a version of (oblivious) Turing Machine, with (potentially) right-infinite tape. Alike standard Turing Machine a rainworm machine is described by the finite set of states $\mathfrak{Q}$, finite tape alphabet $\mathfrak{A}$, finite set of instructions Δ, and an initial configuration.

The "head" of a rainworm machine is always located **not** above one of the cells (like in a usual Turing Machine), but between two consecutive cells (or right to the rightmost cell). So a configuration of an RM can be in a natural way seen as a word from the language $(\mathfrak{A} + \mathfrak{Q})^*$.

The set of states $\mathfrak{Q}$ of a rainworm machine is a disjoint union of $\mathfrak{Q}_0^{\rightarrow}$, $\mathfrak{Q}_0^{\leftarrow}$, $\mathfrak{Q}_1^{\rightarrow}$, $\mathfrak{Q}_1^{\leftarrow}$, of $\mathfrak{Q}_{\gamma_0}^{\rightarrow}$, $\mathfrak{Q}_{\gamma_1}^{\rightarrow}$ and of $\{\eta_{11}, \eta_0, \eta_1\}$.

[14] Neither Figure 3 did, but in the context of **Step 2** we only needed to prove that some pattern **will** appear in the constructed green graph, so – as we were still able to prove it – we could pretend that we did not notice that some of the vertices were pairwise equal.

The finite tape alphabet $\mathfrak{A}$ is a disjoint union of sets $\mathfrak{A}_0$, $\mathfrak{A}_1$ and $\{\alpha, \beta_0, \beta_1, \gamma_0, \gamma_1, \omega_0\}$.

The set Δ of instructions consists of some number of instructions of any of the following forms[15]:

$\diamondsuit_1$: $\eta_{11} \rightsquigarrow \gamma_1\eta_0$

$\diamondsuit_2$: $\eta_0 \rightsquigarrow \mathfrak{b}\eta_1$ where $\mathfrak{b} \in \mathfrak{A}_0$

$\diamondsuit_3$: $\eta_1 \rightsquigarrow \mathfrak{q}\omega_0$ where $\mathfrak{q} \in \mathfrak{Q}_1^{\leftarrow}$

$\diamondsuit_4$: $\mathfrak{b}'\mathfrak{q} \rightsquigarrow \mathfrak{q}'\mathfrak{b}$ where $\mathfrak{q} \in \mathfrak{Q}_0^{\leftarrow}$, $\mathfrak{q}' \in \mathfrak{Q}_1^{\leftarrow}$, $\mathfrak{b} \in \mathfrak{A}_0$ and $\mathfrak{b}' \in \mathfrak{A}_1$
$\diamondsuit_{4'}$: $\mathfrak{b}\mathfrak{q}' \rightsquigarrow \mathfrak{q}\mathfrak{b}'$ where $\mathfrak{q} \in \mathfrak{Q}_0^{\leftarrow}$, $\mathfrak{q}' \in \mathfrak{Q}_1^{\leftarrow}$, $\mathfrak{b} \in \mathfrak{A}_0$ and $\mathfrak{b}' \in \mathfrak{A}_1$

$\diamondsuit_5$: $\gamma_1\mathfrak{q} \rightsquigarrow \beta_1\mathfrak{q}'$ where $\mathfrak{q} \in \mathfrak{Q}_0^{\leftarrow}$, $\mathfrak{q}' \in \mathfrak{Q}_{\gamma_0}^{\rightarrow}$
$\diamondsuit_{5'}$: $\gamma_0\mathfrak{q} \rightsquigarrow \beta_0\mathfrak{q}'$ where $\mathfrak{q} \in \mathfrak{Q}_1^{\leftarrow}$, $\mathfrak{q}' \in \mathfrak{Q}_{\gamma_1}^{\rightarrow}$

$\diamondsuit_6$: $\mathfrak{q}\mathfrak{b} \rightsquigarrow \gamma_1\mathfrak{q}'$ where $\mathfrak{q} \in \mathfrak{Q}_{\gamma_1}^{\rightarrow}$, $\mathfrak{q}' \in \mathfrak{Q}_0^{\rightarrow}$, $\mathfrak{b} \in \mathfrak{A}_0$
$\diamondsuit_{6'}$: $\mathfrak{q}\mathfrak{b} \rightsquigarrow \gamma_0\mathfrak{q}'$ where $\mathfrak{q} \in \mathfrak{Q}_{\gamma_0}^{\rightarrow}$, $\mathfrak{q}' \in \mathfrak{Q}_1^{\rightarrow}$, $\mathfrak{b} \in \mathfrak{A}_1$

$\diamondsuit_7$: $\mathfrak{q}'\mathfrak{b} \rightsquigarrow \mathfrak{b}'\mathfrak{q}$ where $\mathfrak{q} \in \mathfrak{Q}_0^{\rightarrow}$, $\mathfrak{q}' \in \mathfrak{Q}_1^{\rightarrow}$, $\mathfrak{b} \in \mathfrak{A}_0$ and $\mathfrak{b}' \in \mathfrak{A}_1$
$\diamondsuit_{7'}$: $\mathfrak{q}\mathfrak{b}' \rightsquigarrow \mathfrak{b}\mathfrak{q}'$ where $\mathfrak{q} \in \mathfrak{Q}_0^{\rightarrow}$, $\mathfrak{q}' \in \mathfrak{Q}_1^{\rightarrow}$, $\mathfrak{b} \in \mathfrak{A}_0$ and $\mathfrak{b}' \in \mathfrak{A}_1$

$\diamondsuit_8$: $\mathfrak{q}\omega_0 \rightsquigarrow \mathfrak{b}\eta_0$ where $\mathfrak{q} \in \mathfrak{Q}_1^{\rightarrow}$ and $\mathfrak{b} \in \mathfrak{A}_1$

We require the set Δ of instructions to be a partial function[16] – two different instructions must have different left hand sides.

The initial configuration of the machine is $\alpha\eta_{11}$.

As we said before, a configuration of an RM can be in a natural way seen as a word from the language $(\mathfrak{A} + \mathfrak{Q})^*$, so Δ is formulated in the language of Thue semisystem rules[17]. A single computation step of an RM can (should) be seen as a single application of a Thue semi-system rewriting. Following the standard Thue systems notational convention the notation $\mathfrak{w} \rightsquigarrow_\Delta \mathfrak{v}$ means that $\mathfrak{w} = \mathfrak{w}_1\mathfrak{s}\mathfrak{w}_2$ and $\mathfrak{v} = \mathfrak{w}_1\mathfrak{t}\mathfrak{w}_2$ for some rule $\mathfrak{s} \rightsquigarrow \mathfrak{t}$ in Δ. We also use $\overset{k}{\rightsquigarrow}_\Delta$ to denote the k'th power of relation $\rightsquigarrow_\Delta$, $\overset{*}{\rightsquigarrow}_\Delta$ to denote the transitive closure of $\rightsquigarrow_\Delta$ and $\overset{*}{\leftrightsquigarrow}_\Delta$ to denote the symmetric transitive closure of $\rightsquigarrow_\Delta$ (i.e. the smallest equivalence relation having $\rightsquigarrow_\Delta$ as a subset).

While a configuration of a rainworm machine can be always seen as a word, it is clear that not all words from $(\mathfrak{A} + \mathfrak{Q})^*$ make sense as configurations:

Definition 19: Call symbols in $\{\alpha, \beta_0, \gamma_0, \eta_0\} \cup \mathfrak{Q}_0^{\rightarrow} \cup \mathfrak{Q}_0^{\leftarrow} \cup \mathfrak{Q}_{\gamma_0}^{\rightarrow} \cup \mathfrak{A}_0$ even *and symbols in* $\{\beta_1, \gamma_1, \eta_1, \eta_{11}\} \cup \mathfrak{Q}_1^{\rightarrow} \cup \mathfrak{Q}_1^{\leftarrow} \cup \mathfrak{Q}_{\gamma_1}^{\rightarrow} \cup \mathfrak{A}_1$ odd. *A word* $\mathfrak{w} \in (\mathfrak{A} + \mathfrak{Q})^*$ *is an RM configuration, if:*

1) $\mathfrak{w} \in \mathfrak{A}^+\mathfrak{Q}\mathfrak{A}^*$ *(which means there is exactly one symbol in* $\mathfrak{w}$ *which symbolizes the head of the machine);*
2) *the last symbol of* $\mathfrak{w}$ *is one of* η_{11}, η_0, η_1, ω_0*;*
3) *odd and even symbols occur in* $\mathfrak{w}$ *alternately (there are never two odd or two even symbols next to each other);*
4) $\mathfrak{w}$ *is of the form* $\mathfrak{w}_1\mathfrak{w}_2$ *where* $\mathfrak{w}_1$ *is of the form* $\alpha(\beta_1\beta_0)^*$ *or* $\alpha(\beta_1\beta_0)^*\beta_1$, $\mathfrak{w}_2$ *begins with* γ_0 *or* γ_1 *or an element of* $\mathfrak{Q}_{\gamma_0}^{\rightarrow}$ *or an element of* $\mathfrak{Q}_{\gamma_1}^{\rightarrow}$*, and none of* α, β_0, β_1 *occur in* $\mathfrak{w}_2$.

Proof of the following lemma is straightforward case inspection and induction:

Lemma 20: Let $\alpha\eta_{11} \overset{*}{\rightsquigarrow}_\Delta \mathfrak{w}$. *Then* $\mathfrak{w}$ *is an RM configuration.*

Now we are going to explain – informally – how a rainworm creeps. Imagine $\mathfrak{w}$ like in Lemma 20 and let $= \mathfrak{w}_1\mathfrak{w}_2$ be as in Definition 19(4). Suppose the last symbol of $\mathfrak{w}_2$ is η_0. Think of $\mathfrak{w}_2$ as of a rainworm (η_0 being its front and γ_0 or γ_1 being its rear end) and of $\mathfrak{w}_1$ as of the slime trail a rainworm leaves behind. Now there is at most one thing we can do (since Δ is a partial function) – we can use some rule of the form $\diamondsuit_2$ to rewrite η_0 into $\mathfrak{b}\eta_1$ for some $\mathfrak{b}$. Our rainworm has grown one symbol longer! In the next step we can[18] use a rule of the form $\diamondsuit_3$ to rewrite η_1 into $\mathfrak{q}\omega_0$ for some $\mathfrak{q} \in \mathfrak{Q}_1^{\leftarrow}$. The rainworm has grown one symbol longer again, but now the head[19] is no longer in front of the animal. The state of the head is from $\mathfrak{Q}_1^{\leftarrow} \cup \mathfrak{Q}_0^{\leftarrow}$ now. The head will now move, cell by cell, towards the rear end of the rainworm (applying rules $\diamondsuit_4$ and $\diamondsuit_4'$ alternately) rewriting the symbols from $\mathfrak{A}_0 \cup \mathfrak{A}_1$ it passes on its way. Then, after γ_1 (or γ_0) is reached, it is rewritten, by some rule of he form $\diamondsuit_5$ (or $\diamondsuit_5'$) into β_1 (or β_0). The rear end of the rainworm moves towards the front, but the slime trail gets one symbol longer!

Now – since the last rule used was one of $\diamondsuit_5$ (or $\diamondsuit_5'$) – the state is one from $\mathfrak{Q}_{\gamma_0}^{\rightarrow}$ (or $\mathfrak{Q}_{\gamma_1}^{\rightarrow}$), so the next rewriting will move the head to the right and replace the first symbol from $\mathfrak{A}_1$ it encounters with γ_0 (or, resp. the first symbol from $\mathfrak{A}_0$ with γ_1). Only rules of the form $\diamondsuit_7$ or $\diamondsuit_7'$ will be applicable then, and the head will keep moving right (and rewriting the tape symbols on its way), towards the front of the rainworm, until ω_0 is found. Then a rule of the form $\diamondsuit_8$ will be used, and we will be back to the original situation, with η_0 as the rightmost symbol. Notice that the rainworm is longer now (we added one symbol twice and removed one symbol once) and also the slime trail (which is an $\alpha\beta$-path) is longer (we added one symbol). Given Δ, there are of course two possibilities – either the rainworm will creep forever, leaving behind an infinite $\alpha\beta$-slime trail or, at some point, no rule will be applicable, and the process will terminate. It is easy to prove, using textbook techniques, that:

Lemma 21: The problem whether, for given Δ*, the rainworm*[20] *creeps forever, is undecidable.*

B. Creeping back and forth

It is not going to surprise anyone that our next goal is to translate Δ into a set of green graph rewriting rules. But, unlike rainworm machine instructions, green graph rewriting rules are symmetric. This is how we deal with it:

Lemma 22:
1) *If* $\mathfrak{w} \overset{*}{\rightsquigarrow}_\Delta \mathfrak{v}$ *and* $\mathfrak{v}$ *is an RM-configuration then* $\mathfrak{w}$ *satisfies conditions (1)-(3) of Definition 19.*
2) *If* $\mathfrak{w}$ *satisfies condition (1) of Definition 19 then there exists at most one* $\mathfrak{v}$ *such that* $\mathfrak{w} \rightsquigarrow_\Delta \mathfrak{v}$.

[15] Do not give up! An informal explanation will soon come.

[16] In other words, this condition means that rainworm machine is a deterministic computation model.

[17] See e.g. our paper [GM15]. Or [D77] if you prefer a more serious introduction

[18] We should repeat here – and in several other places in this paragraph – that "there is at most one thing we can do now".

[19] Head in the Turing machine sense, rainworms have neither head nor tail. They have front and rear.

[20] Given Δ, the sets $\mathfrak{Q}$ and $\mathfrak{A}$ can be reconstructed.

3) *There is a constant $c_\triangle$, such that if $\mathfrak{v}$ satisfies condition (1) of Definition 19 then there exist at most $c_\triangle$ words $\mathfrak{w}$ such that $\mathfrak{w} \rightsquigarrow_\triangle \mathfrak{v}$.*

Proof: First claim can be proved by straightforward case inspection and induction. The second and the third follow easily from the construction of Δ (remember that Δ is a partial function). □

Now suppose, till the end of this subsection, **that the computation of a rainworm machine with the set of instructions Δ terminates**, after some number $k_\triangle$ of steps, which means that there is $\mathfrak{u}_\triangle$ such that $\alpha\eta_{11} \overset{k_\triangle}{\rightsquigarrow}_\triangle \mathfrak{u}_\triangle$ and no rule from Δ can be applied to $\mathfrak{u}_\triangle$ any more. Then:

Lemma 23: 1) $\{\mathfrak{w} : \mathfrak{w} \overset{*}{\leftsquigarrow}_\triangle \alpha\eta_{11}\} = \{\mathfrak{w} : \mathfrak{w} \overset{*}{\rightsquigarrow}_\triangle \mathfrak{u}_\triangle\}$.

2) *If $\mathfrak{w} \overset{*}{\rightsquigarrow}_\triangle \mathfrak{u}_\triangle$ then $\mathfrak{w}$ satisfies cond. (4) of Definition 19.*

3) *If $\mathfrak{w} \overset{k}{\rightsquigarrow}_\triangle \mathfrak{u}_\triangle$ then $k \leq k_\triangle$.*

4) *The set $\{\mathfrak{w} : \mathfrak{w} \overset{*}{\rightsquigarrow}_\triangle \mathfrak{u}_\triangle\}$ is finite.*

Proof: First claim follows from Lemma 22: to reach any vertex of a tree (and, due to Lemma 22(2) the interesting part of the $\rightsquigarrow_\triangle$-graph is a tree) from a leaf, it is enough to go up to the root and then down. For the proof of the second claim notice that it of course holds for $\mathfrak{u}_\triangle$, and that if it holds for some $\mathfrak{v}_2$, and $\mathfrak{v}_1 \rightsquigarrow_\triangle \mathfrak{v}_2$ then it also holds for $\mathfrak{v}_1$.

To see why the third claim is true notice that it follows from the construction of Δ that, while of course for a configuration $\mathfrak{w}$ of Δ and for a $k \in \mathbb{N}$ there may very well be two different configurations $\mathfrak{v}$ and $\mathfrak{v}'$ such that $\mathfrak{v} \overset{k}{\rightsquigarrow}_\triangle \mathfrak{w}$ and $\mathfrak{v}' \overset{k}{\rightsquigarrow}_\triangle \mathfrak{w}$, always in such case $\mathfrak{v}$ and $\mathfrak{v}'$ will be of equal length and the machine head (that is the symbol from $\mathfrak{Q}$) will be in the same place in $\mathfrak{v}$ and $\mathfrak{v}'$.
Last claim follows from the third and from Lemma 22(3). □

C. From rainworms to green graph rules.

For a rainworm machine Δ we define the set $\mathfrak{T}_\triangle$ of green graph rewriting rules as follows:

- Rules $\emptyset \curlywedge \emptyset \leftrightarrow \alpha \curlywedge \eta_{11}$ and $\eta_{11} \curlyvee \emptyset \leftrightarrow \gamma_1 \curlyvee \eta_0$ are in $\mathfrak{T}_\triangle$;
- $\eta_0 \curlywedge \emptyset \leftrightarrow \mathfrak{b} \curlywedge \eta_1$ is in $\mathfrak{T}_\triangle$ if $\eta_0 \rightsquigarrow \mathfrak{b}\eta_1$ is in Δ;
- $\eta_1 \curlyvee \emptyset \leftrightarrow \mathfrak{q} \curlyvee \omega_0$ is in $\mathfrak{T}_\triangle$ if $\eta_1 \rightsquigarrow \mathfrak{q}\omega_0$ is in Δ;
- $\mathfrak{x} \curlyvee \mathfrak{t} \leftrightarrow \mathfrak{x}' \curlyvee \mathfrak{t}'$ is in $\mathfrak{T}_\triangle$ if $\mathfrak{x}\mathfrak{t} \rightsquigarrow \mathfrak{x}'\mathfrak{t}'$ is an instruction of the form $\diamondsuit_4$, $\diamondsuit_5$, $\diamondsuit_6$, $\diamondsuit_7$ or $\diamondsuit_8$ in Δ;
- $\mathfrak{x} \curlywedge \mathfrak{t} \leftrightarrow \mathfrak{x}' \curlywedge \mathfrak{t}'$ is in $\mathfrak{T}_\triangle$ if $\mathfrak{x}\mathfrak{t} \rightsquigarrow \mathfrak{x}'\mathfrak{t}'$ is an instruction of the form $\diamondsuit'_4$, $\diamondsuit'_5$, $\diamondsuit'_6$ or $\diamondsuit'_7$ in Δ.

Now, **in view of Lemma 12 and Lemma 21, in order to prove Theorem 5 it is enough to show:**

Lemma 24: For given Δ, the rainworm creeps forever (i.e. rainworm machine with Δ as its set of instructions never halts) if and only if the set $\mathfrak{T}^{\boxplus}_\triangle = \mathfrak{T}_\triangle \cup \mathfrak{T}_\boxplus$ of green graph rewriting rules finitely leads to the red spider.

The rest of this Section is devoted to the proof of Lemma 24

D. The "⇒" direction (the easier one).

Suppose Δ is the set of instructions of some rainworm which creeps forever. Then:

Lemma 25:
If $\alpha\eta_{11} \overset{*}{\rightsquigarrow}_\triangle \mathfrak{w}$ then $\mathfrak{w} \in words(chase(\mathfrak{T}_\triangle, \mathbb{D}_\star))$.

Proof: Of course $\alpha\eta_{11} \in words(chase(\mathfrak{T}_\triangle, \mathbb{D}_\star))$, since one gets it from $\mathbb{D}_\star$ by a single application of the first rule of $\mathfrak{T}_\triangle$. For the induction step we are going to show that if $\mathfrak{w} \rightsquigarrow_\triangle \mathfrak{v}$ and $\mathfrak{w} \in words(chase(\mathfrak{T}_\triangle, \mathbb{D}_\star))$ then also $\mathfrak{v} \in words(chase(\mathfrak{T}_\triangle, \mathbb{D}_\star))$.

So suppose $\mathfrak{w} = \mathfrak{w}_1\mathfrak{s}\mathfrak{w}_2$ and $\mathfrak{v} = \mathfrak{w}_1\mathfrak{t}\mathfrak{w}_2$ for a rule $\mathfrak{s} \rightsquigarrow \mathfrak{t}$ in Δ. Let (for example, as all cases are similar) $\mathfrak{s} = \mathfrak{c}_0\mathfrak{c}_1$ and $\mathfrak{t} = \mathfrak{c}'_0\mathfrak{c}'_1$, for some even $\mathfrak{c}_0$ and $\mathfrak{c}'_0$ and odd $\mathfrak{c}_1$ and $\mathfrak{c}'_1$. Then, by construction of $\mathfrak{T}_\triangle$, $\mathfrak{c}_0 \curlywedge \mathfrak{c}_1 \leftrightarrow \mathfrak{c}'_0 \curlywedge \mathfrak{c}'_1$ is a rule of $\mathfrak{T}_\triangle$.

Let c, c' be vertices of the green graph $chase(\mathfrak{T}_\triangle, \mathbb{D}_\star)$ such that: $\mathfrak{w}_1 \in paths(PG(chase(\mathfrak{T}_\triangle, \mathbb{D}_\star)), \mathbf{a}, c)$,
that $\mathfrak{s} \in paths(PG(chase(\mathfrak{T}_\triangle, \mathbb{D}_\star)), c, c')$
and that $\mathfrak{w}_2 \in paths(PG(chase(\mathfrak{T}_\triangle, \mathbb{D}_\star)), c', \mathbf{a})$
(or $\mathfrak{w}_2 \in paths(PG(chase(\mathfrak{T}_\triangle, \mathbb{D}_\star)), c', \mathbf{b})$).

Now, $\mathfrak{s} \in paths(PG(chase(\mathfrak{T}_\triangle, \mathbb{D}_\star)), c, c')$ means that there is a vertex d of $chase(\mathfrak{T}_\triangle, \mathbb{D}_\star)$ such that $H_{\mathfrak{c}_0}(c, d)$ and $H_{\mathfrak{c}_1}(c', d)$ are edges of $chase(\mathfrak{T}_\triangle, \mathbb{D}_\star)$. But, since $chase(\mathfrak{T}_\triangle, \mathbb{D}_\star) \models \mathfrak{T}_\triangle$, the rule $\mathfrak{c}_0 \curlywedge \mathfrak{c}_1 \leftrightarrow \mathfrak{c}'_0 \curlywedge \mathfrak{c}'_1$ is also satisfied in $chase(\mathfrak{T}_\triangle, \mathbb{D}_\star)$, and thus there exists a vertex d' in $chase(\mathfrak{T}_\triangle, \mathbb{D}_\star)$ such that $H_{\mathfrak{c}'_0}(c, d')$ and $H_{\mathfrak{c}'_1}(c', d')$ are edges of $chase(\mathfrak{T}_\triangle, \mathbb{D}_\star)$. This means that $\mathfrak{t} \in paths(PG(chase(\mathfrak{T}_\triangle, \mathbb{D}_\star), c, c')$ and, in consequence, $\mathfrak{v} \in words(chase(\mathfrak{T}_\triangle, \mathbb{D}_\star))$. □

Since rainworm Δ creeps forever, it follows from Lemma 25 that there are $\alpha\beta$-paths of unbounded length in $chase(\mathfrak{T}_\triangle, \mathbb{D}_\star)$, and hence in $chase(\mathfrak{T}^{\boxplus}_\triangle, \mathbb{D}_\star)$. Now we use the machinery from Section VII Step 2 to prove that every finite model of $\mathfrak{T}^{\boxplus}_\triangle$, containing $\mathbb{D}_\star$, contains a 1-2 pattern, so $\mathfrak{T}^{\boxplus}_\triangle$ finitely leads to the red spider.

E. The "⇐" direction (the harder one).

Now we consider a rainworm Δ whose computation terminates. Let $\mathfrak{u}_\triangle = \alpha(\beta_1\beta_0)^n\gamma_1 \dots \omega_0$ be the final configuration[21] like in Section VIII B, and let also $k_\triangle$ be as defined there. Our goal is to construct a finite green graph $\mathfrak{M}$, without a 1-2 pattern, containing $\mathbb{D}_\star$ and being a model of $\mathfrak{T}^{\boxplus}_\triangle$. It would be tempting to think that $chase(\mathfrak{T}^{\boxplus}_\triangle, \mathbb{D}_\star)$ is such a graph. But for some subtle reasons, it is not. Apparently, already $chase(\mathfrak{T}_\triangle, \mathbb{D}_\star)$ can be infinite.

First we will apply a chase-like procedure to construct a finite structure $\mathbb{M}$ being a model of $\mathfrak{T}_\triangle$.

Let $\mathbb{M}^0$ be the structure containing just nothing more than $\mathbb{D}_\star$ and $\mathfrak{u}_\triangle$: there is an edge $H_\emptyset(\mathbf{a}, \mathbf{b})$, then edges $H_\alpha(\mathbf{a}, b_1)$, $H_{\beta_1}(a_1, b_1)$, $H_{\beta_0}(a_1, b_2) \dots H_{\beta_1}(a_n, b_n)$ and so on, and finally there is an edge, labelled with ω_0, from some vertex to $\mathbf{b}$. The structure $\mathbb{M}$ is defined by the following **procedure**:

- **for** m=0 **to** $k_\triangle$ **do**:

{ $\mathbb{M}^{m+1} := \mathbb{M}^m$;
for all pairs c, c' of vertices of $\mathbb{M}^m$ and **all** rules of the form $\mathfrak{c} \curlywedge \mathfrak{d} \leftrightarrow \mathfrak{c}' \curlywedge \mathfrak{d}'$ [or of the form $\mathfrak{c} \curlyvee \mathfrak{d} \leftrightarrow \mathfrak{c}' \curlyvee \mathfrak{d}'$] in $\mathfrak{T}_\triangle$ **if:**

[21] This is only to fix attention and simplify notations. Of course $\mathfrak{u}_\triangle$ does not need to be exactly of this form – the last symbol can also be η_0, η_1 or even η_{11}, and the symbol after last β_0 may very well be different than γ_1.

there exists vertex d' of $\mathbb{M}^m$ such that:
(♠) $\mathbb{M}^m \models H_{\mathfrak{c}'}(c,d'), H_{\mathfrak{d}'}(c',d')$ [or $\mathbb{M}^m \models H_{\mathfrak{c}'}(d',c), H_{\mathfrak{d}'}(d',c')$]
but:
(♡) there is no such vertex d of $\mathbb{M}^m$ that:
$\mathbb{M}^m \models H_{\mathfrak{c}}(c,d), H_{\mathfrak{d}}(c',d)$ [or $\mathbb{M}^m \models H_{\mathfrak{c}}(d,c), H_{\mathfrak{d}}(d,c')$]
do:{

(i) if $\mathfrak{d} \neq \emptyset$, then add to $\mathbb{M}^{m+1}$ a new vertex d and edges $H_{\mathfrak{c}}(c,d)$ and $H_{\mathfrak{d}}(c',d)$ [resp. edges $H_{\mathfrak{c}}(d,c)$, $H_{\mathfrak{d}}(d,c')$];

(ii) if $\mathfrak{d} = \emptyset$, then add to $\mathbb{M}^{m+1}$ an edge $H_{\mathfrak{c}}(c,\mathbf{b})$ [resp. an edge $H_{\mathfrak{c}}(\mathbf{a},c)$];
}};
- $\mathbb{M} := \mathbb{M}^{k_{\triangle}+1}$.

For three reasons this procedure does not (*a priori*) build $chase(\mathfrak{T}_{\triangle}, \mathbb{D}_{\bigstar})$. First, while each rule in $\mathfrak{T}_{\triangle}$ is an equivalence (a conjunction of two TGDs), the procedure executes only one of them, the right-to-left one. Second, in (**ii**), instead of (as *chase* would do) creating a vertex d and two edges $H_{\mathfrak{c}}(c,d)$ and $H_{\emptyset}(\mathbf{a},d)$ it reuses[22] the existing vertex $\mathbf{b}$ and edge $H_{\emptyset}(\mathbf{a},\mathbf{b})$. Third reason is that it terminates after a fixed number of stages, possibly before reaching a fixpoint.

Lemma 26:
- $\mathbb{M}$ *is finite and* $\mathbb{M} \models \mathfrak{T}_{\triangle}$.
- *If* $\mathbb{M} \models H_{\beta_0}(x,y)$, *(or* $\mathbb{M} \models H_{\beta_1}(x,y)$*) for some vertices* x, y, *then* x, y *are already vertices of* $\mathbb{M}^0$ *and* $\mathbb{M}^0 \models H_{\beta_0}(x,y)$, *(or* $\mathbb{M}^0 \models H_{\beta_1}(x,y)$*).*

Proof of Lemma 26 can be found in Appendix C.

Now consider $\mathfrak{M} = chase(\mathfrak{T}^{\boxplus}, \mathbb{M})$. No two different edges labelled with β_0 share the end in $\mathbb{M}$. But, as we saw in Section VII (Step 3) this does not stop the grid triggering rule from being applied. But again, exactly like in Section VII (Step 3) for each $t \leq n$ (where n is the one from $\mathfrak{u}_{\triangle}$ above) nothing more than a harmless grid $\mathbb{M}_t$ will be constructed, not having 1-2-pattern[23]. So $\mathfrak{M} = \mathbb{M} \cup \bigcup_{t \leq n} \mathbb{M}_t$. Since the only edges in $\mathfrak{M}$ which were not in $\mathbb{M}$ are grid edges, we still have $\mathfrak{M} \models \mathfrak{T}_{\triangle}$, so $\mathfrak{M}$ is indeed a finite model of $\mathfrak{T}^{\boxplus}_{\triangle}$, without the 1-2 pattern.

IX. FO non-rewritability. Proof of Theorem 2 (outline).

It is not hard to guess what is going to be our $\mathcal{Q}$ and Q (notations as in Section I-B) – we only know one example of a set of CQs which finitely determines another query but does not determine it. So – let $\mathcal{Q} = Compile(Precompile(\mathfrak{T}))$ and $Q = \exists^* dalt(\bigstar)$. We need to produce, for a given $l \in \mathbb{N}$, two structures, $\mathbb{D}_y$ and $\mathbb{D}_n$, over Σ (we are back to Abstraction Level 0) such that $\mathbb{D}_y$ contains a copy of $dalt(\bigstar)$, $\mathbb{D}_n$ does not contain one, but the views $\mathcal{Q}(\mathbb{D}_y)$ and $\mathcal{Q}(\mathbb{D}_n)$ are not distinguishable by a FO Ehrenfeucht-Fraisse game with l rounds.

Before we go any further we need to understand well enough the sense of $\mathcal{T}_{\mathcal{Q}}$ – the set of green-red TGDs generated by $\mathcal{Q}$. For simplicity, instead of $\mathcal{Q}$, let us concentrate on $\mathcal{Q}_\infty = Compile(Precompile(\mathfrak{T}_\infty))$. It contains six[24] queries:

(IA) $\mathbb{f}_5 \curlywedge \mathbb{f}_6$ (IB) $\mathbb{f}_5^{\alpha} \curlywedge \mathbb{f}_6^{\eta_1}$
(IIA) $\mathbb{f}_7 \curlyvee \mathbb{f}_8^{\eta_1}$ (IIB) $\mathbb{f}_7^{\eta_0} \curlyvee \mathbb{f}_8^{\beta_1}$
(IIIA) $\mathbb{f}_9 \curlywedge \mathbb{f}_{10}^{\eta_0}$ (IIIB) $\mathbb{f}_9^{\eta_1} \curlywedge \mathbb{f}_{10}^{\beta_0}$

We imagine two viewers – Grace and Ruby – each of them is shown her own structure over Σ. Grace will see $\mathbb{D}_y$ and Ruby $\mathbb{D}_n$. The structures are uncolored, but we imagine the one seen by Grace is Green and the one seen by Ruby is Red. Our goal, as we have already said, is to have $\mathbb{D}_y$ with $\bigstar$, $\mathbb{D}_n$ without $\bigstar$, but to make sure that each of the the two girls can see (almost) the same thing. The trick is that they do not see the real structures, but their image under the CQs in $\mathcal{Q}_\infty$.

A. Attempt 1 – $\mathbb{D}_y$ and $\mathbb{D}_n$ as fragments of $chase(\mathcal{T}_{\mathcal{Q}_\infty}, \bigstar)$.

As we said, there must be a copy of the full green spider $\bigstar$ in $\mathbb{D}_y$. Let $\mathbf{a}$ and $\mathbf{b}$ be its tail and antenna. Now, suppose some current versions of $\mathbb{D}_y$ and $\mathbb{D}_n$ are defined. One of the viewers, say Grace, can complain that "Ruby has a tuple $\mathbf{t}$ in $Q(\mathbb{D}_n)$ (for some $Q \in \mathcal{Q}_\infty$) which I do not see in $Q(\mathbb{D}_y)$". And, as we want the two girls to see same, we must add $\mathbf{t}$ to $Q(\mathbb{D}_y)$. But of course we cannot simply add a tuple to $Q(\mathbb{D}_y)$ – we add something to $\mathbb{D}_y$ in order to make sure that $\mathbf{t}$ is in $Q(\mathbb{D}_y)$. And this is exactly what $Q^{R \to G}$ is about.

So, as we see $chase$ (as a procedure), with respect to $\mathcal{T}_{\mathcal{Q}_\infty}$, is all about making sure that the girls see the same. If some tuple t is in $\mathcal{Q}_\infty(dalt(chase_i(\mathcal{T}_{\mathcal{Q}_\infty}, \bigstar)\restriction G))$ then it is also in $\mathcal{Q}_\infty(dalt(chase_{i+1}(\mathcal{T}_{\mathcal{Q}_\infty}, \bigstar)\restriction R))$, and the other way round. And so, **once** $chase$ **reaches a fixpoint, both girls really see the same**. But – and it is an important point – the structures they watch are very much different. Grace watches the daltonisation of $chase(\mathcal{T}_{\mathcal{Q}_\infty}, \bigstar)\restriction G$, and – if we abstract from the Σ-details and see the structure as a swarm – $chase(\mathcal{T}_{\mathcal{Q}_\infty}, \bigstar)\restriction G$ contains edges labelled with $\bigstar, \bigstar^{\alpha}, \bigstar^{\eta_0}, \bigstar^{\eta_1}, \bigstar^{\beta_0}, \bigstar^{\beta_1}$. Ruby watches the daltonisation of $chase(\mathcal{T}_{\mathcal{Q}_\infty}, \bigstar)\restriction R$, and – again seen as a swarm – $chase(\mathcal{T}_{\mathcal{Q}_\infty}, \bigstar)\restriction R$ contains edges labelled with $\bigstar_4, \bigstar_5, \bigstar_6, \bigstar_7, \bigstar_8, \bigstar_9$.

But $chase(\mathcal{T}_{\mathcal{Q}_\infty}, \bigstar)$ is infinite and we need $\mathbb{D}_y$ and $\mathbb{D}_n$ to be finite. On the other hand, we do not really require $\mathcal{Q}_\infty(\mathbb{D}_y)$ and $\mathcal{Q}_\infty(\mathbb{D}_g)$ to be equal. We just need them to be similar enough so that l-rounds Ehrenfeucht-Fraisse game cannot spot the difference. So how about having, as $\mathbb{D}_y$ and $\mathbb{D}_n$, daltonisations of $chase_i(\mathcal{T}_{\mathcal{Q}_\infty}, \bigstar)\restriction G$ and of $chase_i(\mathcal{T}_{\mathcal{Q}_\infty}, \bigstar)\restriction R$? We know from the analysis of Figure 1 and from Remark 10 that at each stage of $chase$ a single execution of a rule from $\mathcal{Q}_\infty$ will be performed: $chase$ begins with an application of rule (IA) (or, strictly speaking, with an application of a TGD generated by (IA), something red will be added then), then (IB), and then rules (IIA), (IIB), (IIIA), (IIIB) are used alternately, building a complicated green-red structure which, after some abstraction, looks like a long green $\alpha\beta$-path (with additional η_0 and η_1 edges leading to $\mathbf{a}$ and $\mathbf{b}$) wrapped with some red edges.

Since there is always only one match in $chase_i(\mathcal{T}_{\mathcal{Q}_\infty}, \bigstar)\restriction G$ of a TGD generated by one of the queries in

[22] It is easy to prove that in the situation of (**ii**) we have $c' = \mathbf{a}$ [or $c' = \mathbf{b}$].

[23] Notice that $\mathbb{M}_t$ has, among others, elements a_i and b_i for all $i \leq t$, and $\mathbb{M}$ also has, among others, elements a_i and b_i for all $i \leq t$. This is not a coincidence – the set union that defines $\mathfrak{M}$ would make no sense otherwise.

[24] Plus the three queries $\mathbb{f}_1^1 \curlywedge \mathbb{f}_2^2, \mathbb{f}_1^3 \curlywedge \mathbb{f}_2^4, \mathbb{f}^3 \curlywedge \mathbb{f}_3^4$, but we do not need to think about them now.

$\mathcal{Q}_\infty$, the structures $\mathcal{Q}_\infty(dalt(chase_i(\mathcal{T}_{\mathcal{Q}_\infty}, \star) \restriction R))$ and $\mathcal{Q}_\infty(dalt(chase_i(\mathcal{T}_{\mathcal{Q}_\infty}, \star) \restriction G))$ will always differ by just one atom. Two long "paths" which differ by just one atom sounds – from the point of view of Ehrenfeucht-Fraisse games – like hope. But unfortunately, it follows from the construction of $\mathcal{Q}_\infty$, that the relation Ruby will see watching $chase_i(\mathcal{T}_{\mathcal{Q}_\infty}, \star) \restriction R$ via query (IIA) is always equal to the relation she will see watching the same structure via (IIB), and that the same holds for queries (IIIA) and (IIIB). But the two equalities never hold simultaneously for relations Grace sees in $chase_i(\mathcal{T}_{\mathcal{Q}_\infty}, \star) \restriction G$. So whatever way we try to prematurely terminate this infinite chase, the single atom of difference will be enough for a FO formula to spot a difference.

B. Attempt 2 – $\mathbb{D}_y$ and $\mathbb{D}_n$ for $\mathcal{Q}_\infty$

Define $chase^L_{2i}(\mathcal{T}_{\mathcal{Q}_\infty}, \star)$ (L standing for "late") as the set of atoms added to $chase_{2i}(\mathcal{T}_{\mathcal{Q}_\infty}, \star)$ at some stage j, where $i \leq j \leq 2i$, together with all elements involved with these atoms (including $\mathbf{a}$ and $\mathbf{b}$). In other words, atoms of $chase^L_{2i}(\mathcal{T}_{\mathcal{Q}_\infty}, \star)$ are atoms of $chase_{2i}(\mathcal{T}_{\mathcal{Q}_\infty}, \star)$ which are not in $chase_i(\mathcal{T}_{\mathcal{Q}_\infty}, \star)$.

As before, structures $\mathcal{Q}_\infty(dalt(chase^L_{2i}(\mathcal{T}_{\mathcal{Q}_\infty}, \star) \restriction G))$ and $\mathcal{Q}_\infty(dalt(chase^L_{2i}(\mathcal{T}_{\mathcal{Q}_\infty}, \star) \restriction R))$ differ, by one atom, at the end of the path. And – unlike $\mathcal{Q}_\infty(dalt(chase_i(\mathcal{T}_{\mathcal{Q}_\infty}, \star) \restriction R))$ and $\mathcal{Q}_\infty(dalt(chase_i(\mathcal{T}_{\mathcal{Q}_\infty}, \star) \restriction G))$, they also differ at the "beginning": their beginnings are exactly as different as the ends of $\mathcal{Q}_\infty(dalt(chase_i(\mathcal{T}_{\mathcal{Q}_\infty}, \star) \restriction R))$ and of $\mathcal{Q}_\infty(dalt(chase_i(\mathcal{T}_{\mathcal{Q}_\infty}, \star) \restriction G))$ are.

Now take i Large Enough (with respect to l) and:

- define $\mathbb{D}_y$ as a disjoint union[25] of $dalt(chase_i(\mathcal{T}_{\mathcal{Q}_\infty}, \star) \restriction G))$; of i copies of $dalt(chase^L_{2i}(\mathcal{T}_{\mathcal{Q}_\infty}, \star) \restriction G))$ and of i copies of $dalt(chase^L_{2i}(\mathcal{T}_{\mathcal{Q}_\infty}, \star) \restriction R))$;
- define $\mathbb{D}_n$ as a disjoint union of $dalt(chase_i(\mathcal{T}_{\mathcal{Q}_\infty}, \star) \restriction R))$; of i copies of $dalt(chase^L_{2i}(\mathcal{T}_{\mathcal{Q}_\infty}, \star) \restriction G))$ and of i copies of $dalt(chase^L_{2i}(\mathcal{T}_{\mathcal{Q}_\infty}, \star) \restriction R))$;

Now an Ehrenfeucht-Fraisse games argument comes, which could not be more standard: each girl can see – via the queries of $\mathcal{Q}_\infty$ – some number of paths ($2i+1$ of them, to be precise). Among the paths Grace can see there is one whose beginning looks the same as the beginning of the Ruby's counterpart of this path (it is $dalt(chase_i(\mathcal{T}_{\mathcal{Q}_\infty}, \star) \restriction G))$), i paths whose beginnings indicate that they come from daltonisation of the green fragment of $chase^L_{2i}(\mathcal{T}_{\mathcal{Q}_\infty}, \star)$ and i paths whose beginnings indicate that they come from daltonisation of the red fragment of $chase^L_{2i}(\mathcal{T}_{\mathcal{Q}_\infty}, \star)$. Notice that exactly the same collection of left ends is seen by Ruby.

Now the path ends: Grace can see $i+1$ ends which indicate that they come from daltonisation of the green fragment of $chase_i(\mathcal{T}_{\mathcal{Q}_\infty}, \star)$ or of $chase^L_{2i}(\mathcal{T}_{\mathcal{Q}_\infty}, \star)$ (the ends of the two kinds paths are identical) and i ends which indicate that hey come from daltonisation of the red fragment of $chase_i(\mathcal{T}_{\mathcal{Q}_\infty}, \star)$ or of $chase^L_{2i}(\mathcal{T}_{\mathcal{Q}_\infty}, \star)$. Since i is Large, the difference between i and $i+1$ is of course not FO-noticeable[26].

[25]Elements $\mathbf{a}$ and $\mathbf{b}$ are seen as constants, so the word "disjoint" does not apply to them – they belong to all the copies.

[26]Saying "FO" we mean FO formula equivalent to some l-round game.

Among the paths Grace can see there is one whose beginning looks the same as the beginning of the Ruby's counterpart of this path and whose end indicates being green. The end of the respective path Ruby can see indicates being red. But since the paths are Long (because i is Large) their beginnings cannot be related to their ends using a FO formula, and in consequence $\mathbb{D}_y$ and $\mathbb{D}_n$ are indeed FO-indistinguishable.

Clearly, there are details that we omitted. Most notably, for Ruby to see anything, via queries (IIA) and (IIIA), in her daltonised copies of $dalt(chase^L_i(\mathcal{T}_{\mathcal{Q}_\infty}, \star) \restriction R))$, two spiders, $dalt(\bigstar_7)$ and $dalt(\bigstar_9)$, both having $\mathbf{a}$ as the tail and $\mathbf{b}$ as the antenna need to be present in $\mathbb{D}_n$. Also, while this outline is stated mainly on the Abstraction Level 1, the real construction must be at Level 0.

C. Last step. Structures $\mathbb{D}_y$ and $\mathbb{D}_n$ for $\mathcal{Q}$.

This was already complicated for $\mathcal{Q}_\infty$, and the construction we really need is for $\mathcal{Q}$. Let $\mathcal{Q}_\boxplus = Compile(Precompile(\mathfrak{T}_\boxplus))$.

Let again i be a Large Enough natural number and imagine the structure $chase(\mathcal{T}_{\mathcal{Q}_\boxplus}, chase_i(\mathcal{T}_{\mathcal{Q}_\infty}, \star))$ (or $chase(\mathcal{T}_{\mathcal{Q}_\boxplus}, chase^L_{2i}(\mathcal{T}_{\mathcal{Q}_\infty}, \star))$). Like in the case of structures $\mathbb{M}_t$ in Section VII, a grid will be constructed, with $chase_i(\mathcal{T}_{\mathcal{Q}_\infty}, \star)$ playing the $\alpha\beta$-path being the southern and eastern border of this grid. But now of course the grid is two-colored: we have our old green grid, intertwining with some red edges, as explained in Remark 10. The structure of this green-red grid is unclear, but the good news is that we do not need to care: since $chase(\mathcal{T}_{\mathcal{Q}_\boxplus}, chase_i(\mathcal{T}_{\mathcal{Q}_\infty}, \star))$ reaches its fixpoint after a finite number of stages (i.e. terminates), we can be sure (as we observed in Section IX-A) that $\mathcal{Q}_\boxplus(dalt(chase(\mathcal{T}_{\mathcal{Q}_\boxplus}, chase_i(\mathcal{T}_{\mathcal{Q}_\infty}, \star)) \restriction G))$ and $\mathcal{Q}_\boxplus(dalt(chase(\mathcal{T}_{\mathcal{Q}_\boxplus}, chase_i(\mathcal{T}_{\mathcal{Q}_\infty}, \star)) \restriction R))$ are simply equal – using the queries in $\mathcal{Q}_\boxplus$ both girls see the same structures there. So the difference between $\mathcal{Q}(dalt(chase(\mathcal{T}_{\mathcal{Q}_\boxplus}, chase_i(\mathcal{T}_{\mathcal{Q}_\infty}, \star)) \restriction G))$ and $\mathcal{Q}(dalt(chase(\mathcal{T}_{\mathcal{Q}_\boxplus}, chase_i(\mathcal{T}_{\mathcal{Q}_\infty}, \star)) \restriction R))$ remains to be one atom – the same one which made $\mathcal{Q}(dalt(chase_i(\mathcal{T}_\mathcal{Q}, \star) \restriction R))$ and $\mathcal{Q}(dalt(chase_i(\mathcal{T}_\mathcal{Q}, \star) \restriction G))$ different (and the difference between $\mathcal{Q}(dalt(chase(\mathcal{T}_{\mathcal{Q}_\boxplus}, chase^L_{2i}(\mathcal{T}_{\mathcal{Q}_\infty}, \star)) \restriction G))$ and $\mathcal{Q}(dalt(chase(\mathcal{T}_{\mathcal{Q}_\boxplus}, chase^L_{2i}(\mathcal{T}_{\mathcal{Q}_\infty}, \star)) \restriction R))$ remains to be two atoms).

Now define $\mathbb{D}_y$ and $\mathbb{D}_n$ like in Section IX-B but instead of each component of the form $dalt(\mathcal{D} \restriction G)$ or $dalt(\mathcal{D} \restriction R)$ of any of the two disjoint unions there (which means – instead of a part of a path) take $dalt(chase(\mathcal{T}_{\mathcal{Q}_\boxplus}, \mathcal{D}) \restriction G)$ and $dalt(chase(\mathcal{T}_{\mathcal{Q}_\boxplus}, \mathcal{D}) \restriction R)$ (which means – take the grid having this part of a path as two of its borders).

Now notice that the crucial argument we needed for FO-indistinguishability in Section IX-B was that the ends of the paths in each of $\mathbb{D}_y$ and $\mathbb{D}_n$ are far enough from each other, so that they cannot be related to each other by FO. The last observation needed to see that (the new) $\mathbb{D}_y$ and $\mathbb{D}_n$ will be FO-indistinguishable is that adding the grid to the path does not decrease the distance between the ends.

X. References

[A11] Foto N. Afrati; *Determinacy and Query Rewriting for Conjunctive Queries and Views*, Theor. Comput. Sci. 412 vol 11; 2011; pages 1005–1021;

[AD98] Serge Abiteboul, Oliver M. Duschka; *Complexity of Answering Queries Using Materialized Views*, Proceedings of 17th ACM SIGACT-SIGMOD-SIGART PODS; 1998; pages 254–263;

[BGO10] Vince Barany, Georg Gottlob and Martin Otto, *Querying the guarded fragment*; Proceedings of ACM/IEEE LICS 2010; pages 1–10;

[CGLV00] Diego Calvanese, Giuseppe De Giacomo, Maurizio Lenzerini, and Moshe Y Vardi, *Answering regular path queries using views*; Proc. of 16th International Conference on Data Engineering, IEEE, 2000; pages 389–398;

[CR97] Chandra Chekuri, Anand Rajaraman, *Conjunctive Query Containment Revisited*, Proceedings of ICDT '97; Springer, Lecture Notes in Computer Science, vol. 1186; 1997; pages 56-70;

[D77] Martin Davis, *Unsolvable Problems*, in em Handbook of Mathematical Logic, J. Barwise ed., North-Holland 1977; pages 567-594;

[DPT99] Alin Deutsch, Lucian Popa and Val Tannen, *Physical Data Independence, Constraints, and Optimization with Universal Plans*, Proceedings of VLDB'99; Morgan Kaufmann; pages 459-470; 1999;

[F15] Nadime Francis, PhD Thesis, 2015;

[FG12] Enrico Franconi, Paolo Guagliardo, *The View Update Problem Revisited*, CoRR, abs/1211.3016, 2012;

[FGZ12] Wenfei Fan, Floris Geerts and Zheng Lixiao, *View Determinacy for Preserving Selected Information in Data Transformations*; Inf. Syst. 37.1; Elsevier 2012; pages 1–12;

[FKN13] Enrico Franconi, Volha Kerhet and Ngo Nhung, *Exact Query Reformulation with First-Order Ontologies and Databases* Journal of Artificial Intelligence Research 48 (2013);

[FSF14] Nadime Francis, Luc Segoufin, Cristina Sirangelo, *Datalog Rewritings of Regular Path Queries using Views*, Proceedings of 17th International Conference on Database Theory (ICDT), Athens, Greece, March 24-28, 2014., 107–118; 2014;

[GM13] Tomasz Gogacz, Jerzy Marcinkowski, *Converging to the Chase–A Tool for Finite Controllability*; Proceedings of ACM/IEEE Symposium on LICS; 2013;

[GM15] Tomasz Gogacz, Jerzy Marcinkowski *The Hunt for a Red Spider: Conjunctive Query Determinacy Is Undecidable* ; ACM/IEEE Symposium on Logic in Computer Science, LICS 2015; pages 281-292;

[H01] Alon Y. Halevy, *Answering Queries Using Views: A Survey*, The VLDB Journal vol 10.4;pages 270–294; 2001;

[JK82] D.S.Johnson, D. S. and A. Klug, *Testing Containment of Conjunctive Queries Under Functional and Inclusion Dependencies*, Proceedings of the 1st ACM SIGACT-SIGMOD Symposium on Principles of Database Systems PODS 82, pages 164–169;

[NSV07] Alan Nash, Luc Segoufin and Victor Vianu; *Determinacy and Rewriting of Conjunctive Queries Using Views: A Progress Report*, ICDT 2007, Springer LNCS 4353; pages 59-73;

[NSV10] Alan Nash, Luc Segoufin and Victor Vianu; *Views and queries: Determinacy and rewriting*, ACM Trans. Database Syst 35; 2010; pages 21:1–21:41

[P11] Daniel Pasailă, *Conjunctive queries determinacy and rewriting*; Proc. ICDT 2011; ACM Press; pages 220-231;

[R06] Ricardo Rosati, *On the decidability and finite controllability of query processing in databases with incomplete information*; Proceedings of ACM PODS 2006; pages 356–365.

XI. Appendix A. Proof of Lemma 12

A. Proof of Claim (1).

This subsection will not be readable without reading (and understanding) Sections IV, V and VI.A of [GM15] first.

We will introduce two operations on structures: $compile$ (turning swarms into structures) and $decompile$ (turning structures into swarms). They will satisfy the following properties:

Lemma 27:

(i) *Let $\mathbb{D}$ be a swarm. If $\mathbb{D} \models \mathcal{T}$ then $compile(\mathbb{D}) \models Compile(\mathcal{T})$. Moreover $\mathbb{D}$ contains a red spider (i.e. contains an atom of the relation $H(\star,\_,\_)$) if and only if $compile(\mathbb{D})$ contains a red spider (i.e. a copy of the full red spider $\star$) and $\mathbb{D}$ contains a green spider (i.e. contains an atom of the relation $H(\star,\_,\_)$) if and only if $compile(\mathbb{D})$ contains a green spider (i.e. a copy of the full red spider $\star$).*

(ii) *Let $\mathbb{D}$ be a relational structure. If $\mathbb{D} \models Compile(\mathcal{T})$ then $decompile(\mathbb{D}) \models \mathcal{T}$. Moreover, $\mathbb{D}$ contains a red spider (i.e. a copy of the full red spider $\star$) if and only if $decompile(\mathbb{D})$ contains a red spider (i.e. contains an atom of the relation $H(\star,\_,\_)$) and $\mathbb{D}$ contains a green spider (i.e. a copy of the full green spider $\star$) if and only if $decompile(\mathbb{D})$ contains a green spider (i.e. contains an atom of the relation $H(\star,\_,\_)$).*

With this Lemma proof of Lemma 12(1) is trivial: Existence of an appropriate model for $\mathcal{T}$ is equivalent to existence of such model for $Compile(\mathcal{T})$.

Definition 28: The swarm $decompile(\mathbb{D})$ is defined as the set of all triples $H(\mathcal{S}, b, c)$ such that $\mathbb{D} \models H(a, b, c)$ and a is the head of a real spider in $\mathbb{D}$ which is isomorphic to $\mathcal{S}$.

This definition is natural. While "$decompiling$" a structure we just abstract from the physical realization of spider's legs. Proof of Lemma 27(ii) is straightforward.

The second construction is a bit more involved and will require some analysis. The reason for this is that we have to figure out how the legs of the spiders should be connected in $compile(\mathbb{D})$: this information is not present in $\mathbb{D}$.

Definition 29: The relational structure $compile(\mathbb{D})$ is defined as follows. Let $\mathbb{D}^0$ be a structure obtained by replacing each edge $H(\mathcal{S}, a, b)$ of $\mathbb{D}$ by a real spider which is isomorphic to $\mathcal{S}$, with b being his antenna and a being his tail. We say that two knees $b_1, b_2 \in \mathbb{D}^0$ (possibly belonging to two different real spiders) are $\sim$-equivalent if and only if the calves connected to them have the same predicate symbol and the same color. We define $compile(\mathbb{D}) = \mathbb{D}^0/\sim$.

In other words Definition 29 says that in order to built $compile(\mathbb{D})$ we create $4s$ additional vertices which are going to serve as calves ($2s$ green and $2s$ red – one for each equivalence class of $\sim$) and connect each head of a spider from $\mathbb{D}$ to appropriate calves.

Clearly, each edge of swarm $\mathbb{D}$ is represented by a real spider in $compile(\mathbb{D})$. We however need to make sure that no new

real spiders emerged, apart from the ones being representations of edges from $\mathbb{D}$. It is important, because such new spiders could serve as arguments for rules in $Complie(\mathcal{T})$, and make the condition that if $\mathbb{D} \models \mathcal{T}$ then $compile(\mathbb{D}) \models Compile(\mathcal{T})$ unsatisfied.

Lemma 30: decompile(compile($\mathbb{D}$))=$\mathbb{D}$

Proof: Let $\mathbb{D}^0$ be like in Definition 29. Each head of a real spider in $\mathbb{D}^0$ is connected to exactly one thigh atom T_j and exactly one thigh atom T^j for each $j \leq s$. Similarly each thigh atom is connected to exactly one calf atom. Thus in $\mathbb{D}^0$ the only real spiders are those which are associated with some edge of $\mathbb{D}$. Therefore $decompile(\mathbb{D}^0) = \mathbb{D}$.

Notice that each knee in $compile(\mathbb{D})$ is connected to exactly one calf: this was true in $\mathbb{D}^0$ and for each $\sim$-class E the calves connected to E are identical[27]. Since no new real spiders emerge in $\mathbb{D}$ compared to $\mathbb{D}^0$, then indeed $decompile(compile(\mathbb{D})) = \mathbb{D}$. □

Proof of Lemma 27 (i): Let $\mathbb{D}$ be a swarm such that $\mathbb{D} \models \mathcal{T}$. We are going to show that $compile(\mathbb{D}) \models Compile(\mathcal{T})$.

Suppose $s = \mathbb{f}_j^i \curlywedge \mathbb{f}_l^k \in Compile(\mathcal{T})$ and let $\mathrm{H}(h_1, a_1, b)$, $\mathrm{H}(h_2, a_2, b)$ be head atoms of two real spiders s_1, s_2 in $compile(\mathbb{D})$, isomorphic to the ideal spiders $\mathbb{\star}^i$,$\mathbb{\star}^k$ respectively (this is of course one of many possible cases, but they are similar). We have to show that there exist two real spiders s_3, s_4 in $compile(\mathbb{D})$ with head atoms of the form $\mathrm{H}(h_3, a_1, b')$, $\mathrm{H}(h_4, a_2, b')$ isomorphic to the spiders $\bigstar_j$,$\bigstar_l$ respectively. Moreover spider s_1 must share calves upper i and lower j with spider s_3 and spider s_2 must share calves upper k and lower l with spider s_4.

Consider edges $H(\mathbb{\star}^i, a_1, b)$ and $H(\mathbb{\star}^k, a_2, b)$ in $\mathbb{D}$. They must exist due to Lemma 30. Since $\mathbb{D} \models \mathbb{f}_j^i \curlywedge \mathbb{f}_l^k$ we know that in $\mathbb{D}$ must exist edges $H(\bigstar_j, a_1, b')$, $H(\bigstar_l, a_2, b')$. Let s_3, s_4 be real spiders in $\mathbb{D}$ associated with those edges. Let s_1^0, s_2^0 s_3^0, s_4^0 be real spiders in $\mathbb{D}^0$ spiders s_1, s_2 s_3, s_4 come from. Since ^{i}-th calves of s_1^0 and s_3^0 are red, the ^{i}-th knees of s_1 and s_3 are $\sim$-equivalent and hence equal in $\mathbb{D}$. The same holds for $_j$-th,^{k}-th, and $_l$-th calves. □

B. Proof of Claim (2)

Proof of Lemma 12(2) has similar high level architecture as proof of Lemma 12(1), but is more complicated.

Let us first remark that it is hard to reason about **arbitrary** finite models of a set of rules $\mathcal{T}$. Much harder than to reason about $chase(\mathcal{T}, \mathcal{D})$ for some $\mathcal{D}$. This is because $chase(\mathcal{T}, \mathcal{D})$ is being built stage by stage, and inductive argument can be very often applied. Minimal models, which we now define, retain some nice properties of $chase$.

Recall that each rule T from $\mathbb{L}_1$ as well as each rule from $\mathbb{L}_2$ postulates, for two edges of the current swarm (or current green graph) satisfying the left-hand side of the rule, existence of *witnesses* – two edges satisfying the right-hand side of the rule.

[27]It is important here, that all those calves share a common end, which is a constant in Σ.

Definition 31: Let $\mathcal{T} \subseteq \mathbb{L}_1$ (resp. $\mathbb{L}_2$) and let $\mathbb{M}$ be a model of $\mathcal{T}$ containing $H(\mathbb{\star}, \mathbf{a}, \mathbf{b})$. The set of important edges of $\mathbb{M}$ is defined inductively. Edge e is important if:

- $e = H(\mathbb{\star}, \mathbf{a}, \mathbf{b})$
- *There exist a rule $T \in \mathcal{T}$ and important edges e_0 and e_1, which satisfy the left-hand side of T, such that e and some other e' in $\mathbb{M}$ are the postulated pair of witnesses.*

We say that $\mathbb{M}$ is a minimal model of $\mathcal{T}$ if $\mathbb{M}$ contains the edge $H(\mathbb{\star}, \mathbf{a}, \mathbf{b})$ and every edge in $\mathbb{M}$ is important.

It is clear that if there exists a model, then there exists a minimal one. One can just take a substructure containing only important edges as a new model. Notice that since importance of edges is defined in an inductive way, we can now use induction to prove our lemmas.

Now – for a fixed set $\mathcal{T}$ of green graph rewriting rules – we define two more operations on structures: $precompile$ (turning green graphs being **minimal models** of $\mathcal{T}$ into swarms) and $deprecomplie$ (turning swarms being **minimal models** of $\mathcal{T}$ into green graphs). They will satisfy the following properties:

Lemma 32:

(i) *Let $\mathbb{D}$ be a swarm without (an edge labelled with) full red spider, a minimal model of $Precompile(\mathcal{T})$. Then $deprecompile(\mathbb{D}) \models \mathcal{T}$. Moreover $deprecompile(\mathbb{D})$ contains no 1-2 pattern.*

(ii) *Let $\mathbb{D}$ be a green graph without 1-2 pattern, a minimal model of $\mathcal{T}$. Then $precompile(\mathbb{D}) \models Precompile(\mathcal{T})$. Moreover $precompile(\mathbb{D})$ contains no (edge labelled with) full red spider.*

Operation *deprecompile*. Let us remind the reader that spider $\mathbb{\star}_J^I$ (or $\bigstar_J^I$) is called *lower* if J is non-empty. In a similar manner:

Definition 33: Rule $\mathbb{f}_{J_1}^{I_1} \curlyvee \mathbb{f}_{J_2}^{I_2}$ (or $\mathbb{f}_{J_1}^{I_1} \curlywedge \mathbb{f}_{J_2}^{I_2}$) from $\mathbb{L}_1$ is lower *if both J_1 and J_2 are non-empty.*

The proof of the following lemma is by (easy) induction, in the spirit of the proof of [GM15, Lemma 16]:

Lemma 34: Let $\mathcal{T} \subseteq \mathbb{L}_1$ be a set of lower rules and let swarm $\mathbb{M}$ be a minimal model of $\mathcal{T}$. Let $\mathcal{S}$ be the spider being the label of some edge in M. Then $\mathcal{S}$ is red if and only if it is lower.

Definition 35: For a swarm $\mathbb{D}$, the green graph $deprecompile(\mathbb{D})$ is the set of full or upper 1-lame green edges from $\mathbb{D}$.

In other words, $deprecompile(\mathbb{D})$ is what remains of swarm $\mathbb{D}$ after removing everything that is not a valid edge of a green graph.

Proof of Lemma 32(i): The green graph $deprecompile(\mathbb{D})$ contains no 1-2 pattern. Otherwise application of the sequence of rules $\mathbb{f}_1^1 \curlywedge \mathbb{f}_2^2$, $\mathbb{f}_1^3 \curlywedge \mathbb{f}_2^4$, $\mathbb{f}^3 \curlywedge \mathbb{f}_3^4$ to this 1-2 pattern would have created a full red spider in $\mathbb{D}$.

Each rule in $\mathcal{T}$ can be seen as a composition of two rules in $Precompile(\mathcal{T})$: if a rule T from $\mathcal{T}$ postulates

existence of some witnesses, then there exists a pair of rules T_1, T_2 in $Precompile(\mathcal{T})$ which implies existence of the same witnesses. Since $\mathbb{D}$ is a model of $Precompile(\mathcal{T})$ it follows that $deprecompile(\mathbb{D})$ is a model of each composition mentioned before. Therefore $deprecompile(\mathbb{D}) \models \mathcal{T}$. □

Operation *precompile*. To transform a green graph into a swarm we have to at least add some red edges, since rules in $\mathbb{L}_1$ require witnesses in color opposite to the color of arguments. As it turns out, it is enough to add red edges produced by a single stage of $chase$:

Definition 36: For a green graph $\mathbb{D}$, a minimal model of $\mathcal{T}$, let $precompile(\mathbb{D})$ be the swarm $chase_1(Precompile(\mathcal{T}), \mathbb{D})$.

In other words, $precompile(\mathbb{D})$ is $\mathbb{D}$ plus all the red edges demanded, as witnesses, by the rules in $Precompile(\mathcal{T})$ with arguments from $\mathbb{D}$. Notice that no green edges are added.

Lemma 37: Suppose $\mathbb{D}$ is a green graph like in Lemma 32(ii). Then no edge in $\mathbb{D}$ is labelled with $\bigstar^3$.

Proof: There is no rule in $\mathcal{T}$ involving $\bigstar^3$, so there is no way for such edge to be important. □

Proof of Lemma 32(ii):

First let us show that there is no edge labelled with $\star$ in $precompile(\mathbb{D})$. The only rule which is not lower in $Precompile(\mathcal{T})$ is $\mathbb{f}^3 \between \mathbb{f}^4_3$, and it is the only one which could produce $\star$. But it would need an edge labelled with $\bigstar^3$ as one of the arguments, and, as we have already noticed, there is no such edge in $\mathbb{D}$.

It remains to show that $precompile(\mathbb{D})$ is indeed a model of $Precompile(\mathcal{T})$.

Clearly, there are needed (red) witnesses for all the rules in $Precompile(\mathcal{T})$ with pairs of green edges as arguments. We need to prove that this is also the case for pairs of red edges as arguments.

Let e_1 and e_2 be two edges in $\mathbb{D}$ sharing antennas (or tails, the argument is analogous), labelled with green spiders $\mathcal{S}_1$ and $\mathcal{S}_2$ and let e'_1 and e'_2, labelled with $\mathcal{S}'_1$ and $\mathcal{S}'_2$ be red edges of $precompile(\mathbb{D})$ added by some rule $\mathbb{f}^{I_1}_{J_1} \between \mathbb{f}^{I_2}_{J_2}$ having e_1 and e_2, as arguments. Edges e'_1 and e'_2 also share antennas, and e_1 shares its tail with e'_1 while e_2 shares its tail with e'_2.

We want to show that whenever e'_1 or e'_2 is an argument of some rule, then the demanded (green) witness was already in $\mathbb{D}$ (and thus is in $precompile(\mathbb{D})$)

There are several cases, none of them difficult. As an example we will consider one of them, analysis is similar for all other cases:

Suppose that none of $\mathcal{S}_1$ and $\mathcal{S}_2$ is $\bigstar$. This means (according to the Rule of Spiders Algebra ♣) that $\mathcal{S}_1 = \bigstar^{I_1}$ and $\mathcal{S}_2 = \bigstar^{I_2}$, with $I_1, I_2 \neq \emptyset$. Notice that $\mathbb{f}^{I_1}_{J_1} \between \mathbb{f}^{I_2}_{J_2}$ is neither $\mathbb{f}^1_2 \between \mathbb{f}^2_2$ (because there is no 1-2 pattern in $\mathbb{D}$) nor $\mathbb{f}^3_1 \between \mathbb{f}^4_2$ (because there is no edge labelled with $\bigstar^3$). Using ♣ again we get that $\mathcal{S}'_1 = \star_{J_1}$ and $\mathcal{S}'_2 = \star_{J_2}$. There are only two rules in $Precompile(\mathcal{T})$ which match with any of $\star_{J_1}$ or $\star_{J_2}$: one of them is $\mathbb{f}^{I_1}_{J_1} \between \mathbb{f}^{I_2}_{J_2}$ and another is some $\mathbb{f}^{I'_1}_{J_1} \between \mathbb{f}^{I'_2}_{J_2}$, with $I_1 \between I_2 \leftrightarrow I'_1 \between I'_2$ being one of the green graph rewriting rules of $\mathcal{T}$. Since $\mathbb{D}$ is a model of $\mathcal{T}$ we know that there are edges e''_1 and e''_2 in $\mathbb{D}$, sharing antennas and such that and e''_1 shares its tail with e'_1 while e''_2 shares its tail with e'_2, with $\bigstar^{I'_1}$ being the label of e''_1 and $\bigstar^{I'_2}$ being the label of e''_2. Now e_1 with e_2 and e''_1 with e''_2 are pairs of witnesses for both possible applications of rules involving e'_1 or e'_2.

Notice that we (silently) used the (easy) observation here, that the joint antenna of e'_1 and e'_2 does not belong to any other edge, and so the only edge that can be used as an argument in a rule of the type "$\between$" together with e'_1 is e'_2 (and the other way round). □

XII. Appendix B. Proof of Lemma 18

Clearly, $\mathbb{M} \models \mathfrak{T}_\infty$, so for the proof of Lemma 18 (ii) we only need to show that $\mathbb{M} \models \mathfrak{T}_\boxplus$.

Call the edges of $\mathbb{M}$ which are labelled with $\bigstar$, $\bigstar^\alpha$, $\bigstar^{\beta_0}$, $\bigstar^{\beta_1}$, $\bigstar^{\eta_0}$, or $\bigstar^{\eta_1}$ *skeleton*, and the edges of $\mathbb{M}$ which are labelled with any of the $\bigstar^{\langle -,-,-,-\rangle}$ *foam*. Among the foam edges we will distinguish between $\bar{b}$-foam and b-foam (depending whether the label is of the form $\bigstar^{\langle -,-,-,\bar{b}\rangle}$ or of the form $\bigstar^{\langle -,-,-,b\rangle}$). We will also refer to north-edges, which are foam edges of the form $\bigstar^{\langle n,-,-,-\rangle}$.

The next lemma follows directly from the construction of $\mathbb{M}$:

Lemma 38:
1) *If e_1 and e_2 are two edges of $\mathbb{M}$ which share at least one vertex, and such that e_1 is a skeleton edge, and e_2 is a foam edge, then e_2 is b-foam.*
2) *If e_1 and e_2 are two edges of $\mathbb{M}$ which share at least one vertex, and such that e_1 is a skeleton edge, and e_2 is a foam edge from some $\mathbb{M}_t$ then e_1 is also an edge of $\mathbb{M}_t$.*
3) *If e and e' are two foam edges of $\mathbb{M}$, e coming from $\mathbb{M}_t$ and e' coming from $\mathbb{M}_{t'}$, for some $t \neq t'$, which share at least one vertex, then they both share this vertex with some skeleton edge (and hence are – due to (1) – are both b-foam).*
4) *If e is a north-edge of some $\mathbb{M}_t$ and e shares its end with some edge e' in $\mathbb{M}$ then e' is also an edge of $\mathbb{M}_t$.*

Now we are ready to prove claim (1) of Lemma 18. Clearly, there is no 1-2 pattern in any of the $\mathbb{M}_t$. So the only way to have it in $\mathbb{M}$ would be if (*) one of the edges of the 1-2 pattern came from $\mathbb{M}_t$ and another from $\mathbb{M}_{t'}$, for some $t \neq t'$. But the two edges of a 1-2 pattern (**) are $\bar{b}$-foam edges which share a vertex. By Lemma 38(3) conditions (*) and (**) are contradictory.

Concerning Lemma 18(2), for each $t \in \mathbb{N}$ there is $\mathbb{M}_t \models \mathfrak{T}_\boxplus$. So again, like in the proof of Lemma 18(1) the only way that $\mathbb{M} \not\models \mathfrak{T}$ could possibly happen would be if we were able to find two edges e and e' which share a vertex, do not come from the same $\mathbb{M}_t$, and such that some rule from $\mathfrak{T}_\boxplus$ applies to them. By Lemma 38(2) none of them is a skeleton edge. By Lemma 38(3) they are both b-foam edges. The only rule of $\mathfrak{T}$ which can rewrite two b-foam edges is the grid triggering rule, but this would require e and e' to share their end-vertices

and one of them to be a north-edge. Which is impossible due to Lemma 38(4). □

XIII. Appendix C. Proof of Lemma 26

By an $\mathbf{ab}$-path in a green graph $\mathcal{D}$ we will mean a directed path in $PG(\mathcal{D})$ leading from $\mathbf{a}$ to $\mathbf{a}$ or to $\mathbf{b}$. By a $\mathfrak{Q}$-edge we will mean an edge labelled with an element of $\mathfrak{Q}$.

A pair $H_{\mathfrak{c}'}(c,d'), H_{\mathfrak{d}'}(c',d')$ [or $H_{\mathfrak{c}'}(d',c), H_{\mathfrak{d}'}(d',c')$] of edges of vertices of a green graph which satisfies condition ♠ for some rule will be called *a right-match*. If also condition ♡ is satisfied for c, c' then we say that they are *an interesting right-match*. Analogously we define *left-match*, and *interesting left-match*.

Notice that a green graph is a model of $\mathfrak{T}_\triangle$ if and only if there are no interesting matches there. Also notice that, at each of its elementary steps[28], the **procedure** which constructs $\mathbb{M}$ finds an interesting right-match and adds a vertex and two edges (or just one edge) which match with the left-hand side of the same rule. Well, it is indeed clear that this is what an elementary step does if $\mathfrak{d} \neq \emptyset$. To see that it is also the case when $\mathfrak{d} = \emptyset$ we need:

Lemma 39:
- *If* $\mathbb{M}^m \models H_\mathfrak{a}(\mathbf{a}, c)$ *then* $\mathfrak{a} \in \{\alpha, \eta_{11}, \eta_1\}$
- *If* $\mathbb{M}^m \models H_\mathfrak{a}(c, \mathbf{b})$ *then* $\mathfrak{a} \in \{\eta_0, \omega_0\}$
- *If* $\mathbb{M}^m \models H_\mathfrak{a}(d, c)$ *and* $\mathfrak{a} \in \{\alpha, \eta_{11}, \eta_1\}$ *then* $d = \mathbf{a}$.
- *If* $\mathbb{M}^m \models H_\mathfrak{a}(c, d)$ *and* $\mathfrak{a} \in \{\eta_0, \omega_0\}$ *then* $d = \mathbf{b}$.

Proof of this Lemma is by straightforward induction/case inspection.

Now one can easily notice that when an elementary step is executed, in the situation when $\mathfrak{d} = \emptyset$, then $H_\emptyset(\mathbf{a}, \mathbf{b})$, together with the newly added edge, form the demanded left-match.

Since neither β_0 nor β_1 ever occur in the left-hand side, the second claim of Lemma 26 is obvious. The first claim may seem to be less straightforward. If we only use the rules of $\mathfrak{T}_\triangle$ in the right-to-left direction, how can we be sure that the resulting green graph has no interesting left-matches? And also, if we terminate the **procedure** after a fixed number ($k_\triangle$) of steps how can we be sure there are no interesting right-matches left?

It is straightforward to prove by induction that for each $0 \leq m \leq k_\triangle + 1$:

Lemma 40 (loop invariants):
1) *If* $\mathfrak{v} \in words(\mathbb{M}^m)$ *then* $\mathfrak{v} \overset{*}{\rightsquigarrow}_\triangle \mathfrak{u}_\triangle$.
2) *Every path in* $\mathbb{M}^m$ *is a subpath of some* $\mathbf{ab}$*-path in* $PG(\mathbb{M}^m)$. *In particular, every edge of* $\mathbb{M}^m$ *belongs to some* $\mathbf{ab}$*-path in* $PG(\mathbb{M}^m)$.
3) *On each* $\mathbf{ab}$*-path in* $PG(\mathbb{M}^m)$ *there is exactly one* $\mathfrak{Q}$*-edge.*
4) *Every* $\mathfrak{Q}$*-edge in* $PG(\mathbb{M}^m)$ *is on exactly one* $\mathbf{ab}$*-path in* $PG(\mathbb{M}^m)$.

For the induction step it is crucial to realize that if a rule of $\mathfrak{T}_\triangle$ is of the form $\mathfrak{c}\curlyvee\mathfrak{d} \leftrightarrow \mathfrak{c}'\curlyvee\mathfrak{d}'$ then $\mathfrak{c}$ and $\mathfrak{c}'$ are odd and $\mathfrak{d}$ and $\mathfrak{d}'$ are even and if some rule of $\mathfrak{T}_\triangle$ is of the form $\mathfrak{c}$ $\curlywedge\mathfrak{d} \leftrightarrow \mathfrak{c}'\curlywedge\mathfrak{d}'$ then $\mathfrak{c}$ and $\mathfrak{c}'$ are even and $\mathfrak{d}$ and $\mathfrak{d}'$ are odd. In consequence, in one elementary step of the **procedure** a path of length 2 (or 1), from some c to c' is added in $PG(\mathbb{M}^{m+1})$ only if already in $PG(\mathbb{M}^m)$ there already was a path of length 2 from c to c'

Since – as it follows from Lemma 40(4) – each $\mathfrak{Q}$-edge e of $PG(\mathbb{M}^m)$ is on exactly one $\mathbf{ab}$-path in $PG(\mathbb{M}^m)$, and we know that this path can be read as a word, we can call this word $\mathfrak{w}(e)$.

Lemma 41: There are no interesting left-matches in $\mathbb{M}^m$

Proof: This is of course true in $\mathbb{M}^0$, since $\mathfrak{u}_\triangle$ is the configuration in which Δ terminates. For the induction step imagine a vertex and two edges e and e' (or just one edge, the argument is analogous) were added to the current structure, as the result of one of the elementary steps. It follows from the form of the rules of $\mathfrak{T}_\triangle$ that one of them – let it be e – is a $\mathfrak{Q}$-edge. Rainworms are deterministic, so there exists at most one rule of Δ which can be applied to $\mathfrak{w}(e)$, and thus there is at most one rule in $\mathfrak{T}_\triangle$ whose left-hand side matches with some edges on the $\mathbf{ab}$-path the edge e belongs to. But we already know one such rule – it is the one that created e and e', which means that the edges that match with its right-hand side must have existed in the structure before e and e' were created. □

Lemma 42: For each $m \leq k_\triangle$ *the structure* $\mathbb{M}^m$ *is finite.*

Proof: $\mathbb{M}^0$ is finite. If $\mathbb{M}^m$ is finite then there is a finite number of $\mathfrak{Q}$-edges there, and for each of them there are at most $c_\triangle$ right-matches, so $\mathbb{M}^{m+1}$ is also finite. □

It follows from Lemma 40(1) that if e is a $\mathfrak{Q}$-edge of $\mathfrak{M}$ then $\mathfrak{w}(e) \overset{*}{\rightsquigarrow}_\triangle \mathfrak{u}_\triangle$ and, in consequence (from Lemma 23(1) and (3)), we have $\mathfrak{w}(e) \overset{k_e}{\rightsquigarrow}_\triangle \mathfrak{u}_\triangle$ for some natural number $k_e \leq k_\triangle$.

On the other hand, there of course exists, for each edge e in $PG(\mathbb{M})$, the minimal number m such that e is already in $PG(\mathbb{M}^m)$. This minimal m will be denoted as m_e. Last lemma we need for the proof of Lemma 26 is:

Lemma 43: Let e *be a* $\mathfrak{Q}$*-edge of* $PG(\mathbb{M}^m)$. *Then* $m_e = k_e$.

It follows from the form of the rules in $\mathfrak{T}_\triangle$ that each edge of $\mathbb{M}$ must have been created together with some $\mathfrak{Q}$-edge. So once Lemma 43 is proven we know that $\mathbb{M} = \mathbb{M}^{k_\triangle+1} = M^{k_\triangle}$, which means that the last execution of the main loop in the procedure defining $\mathbb{M}$ was in vain – there were no interesting right-matches in $M^{k_\triangle}$ any more. This – together with Lemma 41 and Lemma 42 will imply that $\mathbb{M} \models \mathfrak{T}_\triangle$.

Proof of Lemma 43: The claim is of course true for $m = 0$.

For the induction step first notice that if edges e and e' are a right-match in $\mathbb{M}^m$ and $m' > m$ then e and e' do not form an interesting right-match in $\mathbb{M}^{m'}$ – this is since condition ♡ was no longer satisfied for them in $\mathbb{M}^{m+1}$ at latest. Notice also that if edges e and e' are a right-match in $\mathbb{M}$ and e is a $\mathfrak{Q}$-edge then e' exists already in $\mathbb{M}^{m_e}$ (this is because e only belongs to one $\mathbf{ab}$-path in $\mathbb{M}$ and this path exists already in $\mathbb{M}^{m_e}$ since we know e belongs to some $\mathbf{ab}$-path already

[28] By "elementary step" we mean a single execution of the inner loop.

in $\mathbb{M}^{m_e}$). In consequence, a $\mathfrak{Q}$-edge e can be a part of an interesting right-match only in $\mathbb{M}^{m_e}$.

Assume now the claim is true for some m and let e_1 be a $\mathfrak{Q}$-edge such that $m_{e_1} = m + 1$. This means there was an interesting right match in $\mathbb{M}^m$, consisting of some e and e', with e being a $\mathfrak{Q}$-edge, such that e_1 was created (possibly together with some e'_1) in some elementary step using this right match. Of course (this is how the rules of $\mathfrak{T}_{\triangle}$ work) we have $\mathfrak{w}(e_1) \rightsquigarrow_{\triangle} \mathfrak{w}(e)$, so $k_{e_1} = k_e + 1$. From the reasoning in the previous paragraph we get that $m = m_e$ and by hypothesis, we have that $m_e = k_e$. So $k_{e_1} = m + 1 = m_{e_1}$. $\square$